

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 ; Search time 27.8333 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDPRSEEDKRFESHIECR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	20	AAW42921	Immunogenic Hepati
2	109	100.0	21	AAW42921	Synthetic HAV VPI
3	109	100.0	854	AAW50287	Sequence encoded b
4	109	100.0	1077	AAW5559	A partial hepatiti
5	109	100.0	1091	AAW32426	Translated from 5'
6	109	100.0	2227	AAW05697	Attenuated hepatit
7	109	100.0	2227	AAW34074	Hepatitis A virus
8	109	100.0	2227	AAW18607	Amino acid sequenc
9	109	100.0	2227	AAW18608	Amino acid sequenc
10	109	100.0	2227	AAW18609	Amino acid sequenc

11	106	97.2	366	6	AAW50230	Sequence of hepati
12	106	97.2	993	6	AAW50116	Sequence of Hepati
13	106	97.2	993	6	AAW50231	Sequence encoded b
14	106	97.2	2227	23	AAE19899	Hepatitis A virus
15	87	79.8	2227	7	AAW60066	Sequence of viral
16	81	74.3	839	12	AAW15629	Capsid region of c
17	69	63.3	20	18	AAW42920	Immunogenic Hepati
18	69	63.3	21	22	AAW69437	Synthetic HAV VPI
19	69	63.3	300	10	AAW50612	Portion of structu
20	46.5	42.7	183	21	AAW57588	Arabidopsis thalia
21	46.5	42.7	187	21	AAW57587	Arabidopsis thalia
22	46.5	42.7	200	21	AAW57586	Arabidopsis thalia
23	46	42.2	306	22	ABG22935	Novel human diagno
24	46	42.2	701	22	ABG01389	Novel human diagno
25	44.5	40.8	1334	20	AAW02630	Protein encoded by
26	44	40.4	131	21	AAW47694	Arabidopsis thalia
27	44	40.4	134	21	AAW47693	Arabidopsis thalia
28	44	40.4	231	20	AAW37694	Chlamydia trachoma
29	44	40.4	285	22	AAW78648	Human protein SEQ
30	44	40.4	292	22	AAW79632	Human protein SEQ
31	44	40.4	326	21	AAW50172	Arabidopsis thalia
32	44	40.4	383	21	AAW50171	Arabidopsis thalia
33	44	40.4	498	21	AAW50170	Arabidopsis thalia
34	44	40.4	598	23	AAU72997	Neisseria meningit
35	43	39.4	125	21	AAW01681	Human secreted pro
36	43	39.4	125	22	AAU23005	Novel human enzyme
37	43	39.4	292	21	AAW26308	Arabidopsis thalia
38	43	39.4	293	21	AAW26307	Arabidopsis thalia
39	43	39.4	322	21	AAW26987	Arabidopsis thalia
40	43	39.4	325	21	AAW32082	Arabidopsis thalia
41	43	39.4	325	21	AAW50197	Arabidopsis thalia
42	43	39.4	379	21	AAW26986	Arabidopsis thalia
43	43	39.4	382	21	AAW32081	Arabidopsis thalia
44	43	39.4	382	21	AAW50196	Arabidopsis thalia
45	43	39.4	429	21	AAW42049	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW42921
ID AAW42921 standard; peptide; 20 AA.
XX
XX AAW42921;
AC
XX
DT 28-APR-1998 (first entry)
XX
XX Immunogenic Hepatitis A virus peptide YK-1314.
DE
DE Immunogenic peptide; immunogenic epitope; VPI protein;
KW
KW immune response; antibody.
XX
XX Synthetic.
OS
OS Hepatitis A virus.
XX
XX W09740147-AL.
PN
XX
PD 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06891.
PF
XX
XX 19-APR-1996; 96US-0015644.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
FA
XX
XX Fields HA, Khudyakov YE;
PI
XX
XX WPI; 1997-535831/49.
DR
XX
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX Claim 13; Page 111; 140pp; English.

XX Peptides AAW42906-21 are immunogenic peptides corresponding to

CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are

CC substantially similar to a portion of the amino acid sequence of the VP1

CC protein of HAV corresponding to amino acids 492-791. The present peptide

CC is derived from amino acids 779-798, and has a reactivity of 54.2% with

CC acute sera. Compositions containing the peptides can be used to induce an

CC immune response to HAV in a mammal. The peptides can also be used to

CC detect the presence of antibodies against HAV in mammalian serum. The

CC peptides can also be used to make an antibody against HAV by

CC administering the peptide to a mammal.

XX Sequence 20 AA;

Query Match 100.0%; Score 109; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.2e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20

|||||

Db 1 VDDPRSEEDKRFESHIECRK 20

RESULT 2

AAW69438

ID AAW69438 standard; Peptide; 21 AA.

XX AC AAW69438;

XX 20-APR-2001 (first entry)

XX Synthetic HAV VP1 peptide, SEQ ID NO: 38.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;

KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX 25-JAN-2001.

XX 14-JUL-2000; 2000WO-US19267.

XX 15-JUL-1999; 99US-0144412.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays

PT detecting anti-hepatitis A virus and as vaccines -

XX Claim 10; Page 92; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are

CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides

CC comprise antigenic epitopes of the major structural capsid polypeptides

CC or non-structural polypeptides of HAV with one or more glutamine

CC molecules at the carboxy end of the peptide. The peptides are used to

CC detect the presence of antibodies against HAV in mammalian serum, to

CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting

CC IgM antibodies in mammalian serum and detecting convalescence in a

CC mammal. The peptides are used to detect or quantify HAV antibodies in

CC samples in clinical or research-based assays using immunoblotting,

CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy

CC end of the peptides enhances the IgM antibody reactivity.

XX Sequence 21 AA;

Query Match 100.0%; Score 109; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 6.6e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20

|||||

Db 1 VDDPRSEEDKRFESHIECRK 20

RESULT 3

AAW50287

ID AAW50287 standard; Protein; 854 AA.

XX AC AAW50287;

XX 30-NOV-1991 (first entry)

XX Sequence encoded by hepatitis A virus (HAV) cDNA from near the

DE genome 5' terminus to the end of the area corresponding to the

DE capsid protein region of poliovirus RNA.

XX Hepatitis A virus assay; antigen; antibody.

XX Hepatitis A virus.

XX WO8501517-A.

XX 11-APR-1985.

XX 27-SEP-1984; 84WO-US01552.

XX 30-SEP-1983; 83US-0537911.

XX (MASI) MASSACHUSETTS INST TECH.

XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;

PI Racaniello VR;

XX WPI; 1985-098846/16.

DR N-PSDB; AAN50330.

XX New hepatitis A virus cDNA - useful in assays for the virus and

PT for prodn. of the viral antigen and antibodies to it

XX Example; Fig 7; 60pp; English.

XX The inventors claim HAV cDNA and a method for producing it, whereby

CC large amts. can be obtd. economically. The cDNA is useful in the

CC assay for detection of HAV quickly and easily and with high

CC sensitivity and specificity. The HAV cDNA is also used in the prodn.

CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.

XX Sequence 854 AA;

Query Match 100.0%; Score 109; DB 6; Length 854;

Best Local Similarity 100.0%; Pred. No. 3.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20

|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4

AAW95559

ID AAW95559 standard; Protein; 1077 AA.

XX

AC AAW95559;
XX
DT 28-APR-1999 (first entry)
XX
DE A partial hepatitis A virus (HAV) protein.
XX
KW Hepatitis A virus protein; HAV; P2 region;
KW cell-culture-adapted HAV strain; infection; accelerated growth.
XX
OS Hepatitis A virus.
XX
PN US5849562-A.
XX
PD 15-DEC-1998.
XX
PF 06-JUN-1995; 95US-0468926.
XX
PR 06-NOV-1991; 91US-0788262.
PR 30-SEP-1983; 83US-0537911.
PR 27-SEP-1984; 84US-0654942.
PR 06-OCT-1988; 88US-0256135.
PR 06-JUN-1995; 95US-0468926.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Emerson SU, Purcell RH;
XX
DR WPI: 1999-094412/08.
DR N-PSDB; AAX01006.
XX
PT Chimeric hepatitis A virus strains - with P2 region from
PT cell-culture-adapted strain in wild-type genome
XX
PS Disclosure; Fig 7A-L; 36pp; English.
XX
CC The present sequence represents a partial hepatitis A virus (HAV)
CC protein. The specification describes a DNA construct consisting
CC of a wild-type HAV genome in which the P2 region is replaced by the
CC P2 region from a cell-culture-adapted HAV strain. The construct is
CC used to demonstrate that mutations in the P2 region of a
CC cell-culture-adapted HAV strain are sufficient for establishment of
CC infection and accelerated growth in cell culture.
XX
SQ Sequence 1077 AA;
Query Match 100.0%; Score 109; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 1 VDDPRSEEDKRFESHIECRK 20
|||||
DB 1002 VDDPRSEEDKRFESHIECRK 1021
RESULT 5
AAR32426
ID AAR32426 standard; Protein; 1091 AA.
XX
AC AAR32426;
XX
DT 17-DEC-2001 (updated)
DT 10-JUN-1993 (first entry)
XX
DE Translated from 5' region of Hepatitis A Virus genomic clone.
XX
KW HAV HM-175; chronic liver disease; picornavirus.
XX
OS Hepatitis A Virus.
XX
FH Key Location/Qualifiers
FH 238..1091
FT Region /label= ORF
FT /note= "second putative initiation codon at

FT Region 1..711 position 240"
FT
FT /note= "X's correspond to nonsense codons,
XX i.e. this region is not an ORF"
PN USN788262-N.
XX
PD 15-DEC-1992.
XX
PF 30-SEP-1983; 83US-0536911.
XX
PR 27-SEP-1984; 84US-0654942.
PR 06-OCT-1988; 88US-0256135.
PR 30-SEP-1983; 83US-0536911.
PR 06-NOV-1991; 91US-0788262.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Baltimore D, Feinstein SM;
PI Purcell RH, Racanelli VR, Ticehurst JR;
XX
DR WPI: 1993-067429/08.
DR N-PSDB; AAQ36934.
XX
PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
PT of antigen and antibodies
XX
PS Disclosure; Fig 7; 65pp; English.
XX
CC HAV virion RNA was extracted from the livers of marmosets which had
CC been inoculated with HAV (the HAV had previously been passaged twice
CC in marmosets). The RNA was used to prepare ds cDNA clones by
CC standard methods. Clones contg. inserts which hybridized to RNA from
CC HAV-infected African Green Monkey Kidney cells were selected for
CC further analysis. A 7.4kb restriction map (about 99% of the HAV
CC genome) was constructed from 5 overlapping inserts. The sequence of
CC the first 3.3kb (approx.) from the 5'-terminus was determined. An
CC amino acid sequence was decoded from the entire clone and an open
CC reading frame was identified starting at position 238. A comparison
CC of the predicted HAV amino acid sequences with the known capsid
CC protein sequences of other picornaviruses (poliovirus, foot and
CC mouth disease virus and encephalomyelitis virus) revealed areas of
CC local homology.
CC (Note: Revised entry submitted to correct the patent number format of
CC US Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis_us.html.)
XX
SQ Sequence 1091 AA;
Query Match 100.0%; Score 109; DB 14; Length 1091;
Best Local Similarity 100.0%; Pred. No. 4.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
DB 1016 VDDPRSEEDKRFESHIECRK 1035
RESULT 6
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX
AC AAR05697;
XX
DT 15-AUG-1990 (first entry)
XX
DE Attenuated hepatitis A virus.
XX
KW Hepatitis A virus; vaccine; attenuated.
XX
OS Hepatitis A virus, strain HM-175.
XX

Key Location/Qualifiers
FH Region 1..23
FT /label=VP4 = 1A
FT Region 24..245
FT /label=VP2 = 1B
FT Region 246..491
FT /label=VP3 = 1C
FT Region 492..791
FT /label=VP1 = 1D
FT Region 792..980
FT /label=2A
FT Region 981..1087
FT /label=2B
FT Region 1088..1422
FT /label=2C
FT Region 1423..1496
FT /label=3A
FT Region 1497..1519
FT /label=3B = VPg
FT Region 1520..1738
FT /label=3C
FT Region 1739..2227
FT /label=3D

US4894228-A.

16-JAN-1990.

12-JUL-1988; 88US-0217824.

12-JUL-1988; 88US-0217824.

12-JUL-1988; 88US-0652967.

(USSH) US DEPT HEALTH & HUMAN.

Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM;

Daemer RJ, Gust ID;

WPI; 1990-075557/10.

N-PSDB; AAQ03512.

Vaccine against hepatitis A virus infection - comprises novel
attenuated hepatitis A virus strain.

Claim 1; Fig 1; 18pp; English.

The attenuated HAV is useful for inducing protective immunity against
HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
several nucleotide changes distributed throughout the genome, is
attenuated for chimpanzees, elicits serum neutralising antibodies, and is
suitable for use as an HAV vaccine. It is noted that not all the changes
are necessary for attenuation and use as a vaccine.

Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20

|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 7

AAW34074

ID AAW34074 standard; Protein; 2227 AA.

AAW34074;

27-APR-1998 (first entry)

Hepatitis A virus HM-175 protein sequence.

XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW infection; vaccine.
XX Hepatitis A virus HM-175.
XX Key Location/Qualifiers
FH Protein 1..23
FT /label= VP4
FT Protein 24..245
FT /label= VP2
FT Protein 246..491
FT /label= VP3
FT Protein 492..791
FT /label= VP1
FT Protein 792..980
FT /label= 2A
FT Protein 981..1087
FT /label= 2B
FT Protein 1088..1422
FT /label= 2C
FT Protein 1423..1496
FT /label= 3A
FT Protein 1497..1519
FT /label= 3B
FT Protein 1520..1738
FT /label= 3C
FT Protein 1739..2227
FT /label= 3D

W09740166-A2.

30-OCT-1997.

18-APR-1997; 97WO-US06506.

19-APR-1996; 96US-0015642.

(USSH) US SEC DEPT HEALTH.
(USSH) US DEPT HEALTH & HUMAN SERVICES.

Emerson SU, Purcell RH, Raychaudhuri G;

WPI; 1997-535850/49.

N-PSDB; AAT93023.

Human attenuated HAV genome containing simian HAV 2C gene - useful
as vaccines against HAV infection

Disclosure; Fig 13A-D; 66pp; English.

This protein sequence is encoded by the human hepatitis A virus
(HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
HAV/7 is obtained by passage of HM-175 in African Green Monkey
kidney cells. A claimed DNA construct (1) comprises a genome of
HAV, where the genome is a human attenuated HAV genome in which a
region of the 2C gene has been replaced by a corresponding region
from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
region of the 2C gene from AGM-27 contained in the construct
preferably encodes amino acids 120-328 of the 2C protein, amino
acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
transcript of (1); (2) a cell transfected with (1) or the RNA
transcript of (1); (3) a HAV genome as above; (4) antibodies to the
HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
its RNA transcript, can be used as a vaccine for preventing HAV in
a mammal. (1) or the RNA transcript can also be used to stimulate
the production of protective antibodies in the mammal.

Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 18; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 779 VDDPRSEEDKRFESHIECRK 798

RESULT 8
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
AC AAB18607;
XX
DT 15-JAN-2001 (first entry)
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
DR WPI; 2000-586464/55.
DR N-PSDB; AAA75476.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
PS Disclosure; Fig 6A-K; 72pp; English.
XX
CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 779 VDDPRSEEDKRFESHIECRK 798

RESULT 9
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
AC AAB18608;
XX
DT 15-JAN-2001 (first entry)
DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW P-35 virus.

```

```

XX Hepatitis A virus.
OS US6113912-A.
PN
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI; 2000-586464/55.
DR N-PSDB; AAA75477.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
PS Disclosure; Columns 67-78; 72pp; English.
XX
CC The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 779 VDDPRSEEDKRFESHIECRK 798

RESULT 10
AAB18609
ID AAB18609 standard; Protein; 2227 AA.
XX
AC AAB18609;
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW HAV 4380.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

```

XX
DR WPI: 2000-586464/55.
DR N-PSDB; AAA75478.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
XX Disclosure; Columns 93-104; 72pp; English.
PS
PS The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.
XX
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 109; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-07; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECRK 20
Db 779 VDDPRSEEDKRFESHIECRK 798
RESULT 11
AAP50230
ID AAP50230 standard; Protein; 366 AA.
XX
AC AAP50230;
XX
XX 28-NOV-1991 (first entry)
DT
XX
XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).
DE
XX
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
KW
XX
XX Hepatitis A virus.
OS
XX
PN EP138704-A.
XX
XX 24-APR-1985.
PD
XX
XX 09-OCT-1984; 84EP-0402025.
XX
PR 02-MAR-1984; 84US-0585942.
PR 14-OCT-1983; 83US-0541836.
XX
XX (MERI) MERCK & CO INC.
PA
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
PI
XX
XX WPI: 1985-100818/17.
DR N-PSDB; AAN50274.
DR
XX
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus
PT
XX
PS Claim 21; Page 46-48; 49pp; English.
XX
XX VPI is isolated by solubilisation of the intact virus in an aq.
CC anionic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.
CC
XX
XX Sequence 366 AA;
SQ
Query Match 97.2%; Score 106; DB 6; Length 366;
Best Local Similarity 95.0%; Pred. No. 4.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECRK 20
Db 288 VDDPRSEEDKRFESHIECRK 307
RESULT 12
AAP50116
ID AAP50116 standard; Protein; 993 AA.
XX
XX AC AAP50116;
XX
XX 30-SEP-1991 (first entry)
DT
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides
DE VP-1, VP-2, VP-3 and VP-4.
DE
XX
XX Antigenic protein; immunogen; vaccine.
KW
XX
XX Hepatitis A virus (strain CR326).
OS
XX
PN EP154587-A.
XX
PD 11-SEP-1985.
XX
PF 27-FEB-1985; 85EP-0400369.
XX
PR 02-MAR-1984; 84US-0585818.
XX
XX (MERI) MERCK & CO INC.
PA
XX
XX Linemeyer DL, Menke JG, Reuben RG, Mitra SW;
PI
XX
XX WPI: 1985-224964/37.
DR N-PSDB; AAN50139.
DR
XX
XX New nucleotide sequences coding for hepatitis A virus antigens -
PT useful for eliciting normal immune response and in vaccines for
PT protecting against the virus
PT
XX
PS Example; Page 11-17; 32pp; English.
XX
XX Within the sequence in AAN50139 is encoded the information necessary
CC to make the antigenic proteins of HAV. The sequences encoding for
CC the structural proteins begin at base 403. The key sub-unit
CC sequences within VP-1, designated Sequences I, II, III, IV, and V,
CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
CC nucleotide sequences which are valuable as encoding antigenic
CC proteins are the sequences from base 1749 to base 2722; from base
CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
CC base 1749 to base 2722 is esp. valuable as a vector for producing
CC antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
CC translation of a stop codon.
XX
XX Sequence 993 AA;
SQ
Query Match 97.2%; Score 106; DB 6; Length 993;
Best Local Similarity 95.0%; Pred. No. 1.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECRK 20
Db 915 VDDPRSEEDKRFESHIECRK 934
RESULT 13
AAP50231
ID AAP50231 standard; Protein; 993 AA.
XX
XX AC AAP50231;
XX
XX 28-NOV-1991 (first entry)
DT

XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
OS Hepatitis A virus.
XX
FH Key Location/Qualifiers
FT Protein 628..993
FT /note= "claimed; x denotes translated stop codons
FT and unspecified triplets"
XX
XX EP138704-A.
XX
XX 24-APR-1985.
XX
XX 09-OCT-1984; 84EP-0402025.
XX
XX 02-MAR-1984; 84US-0585942.
XX
XX 14-OCT-1983; 83US-0541836.
XX
XX (MERI) MERCK & CO INC.
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
XX
XX WPI; 1985-100818/17.
XX
XX N-PSDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
XX neutralising antibodies to the virus
XX
XX Disclosure; Page 17-23; 49pp; English.
XX
XX VP1 is isolated by solubilisation of the intact virus in an aq.
XX anionic surfactant and a reducing agent. The viral proteins are sepd.
XX and the protein of molecular wt. 33000 daltons is sepd.
XX
XX Sequence 993 AA;
XX
XX Query Match 97.2%; Score 106; DB 6; Length 993;
XX Best Local Similarity 95.0%; Pred. No. 1.2e-07;
XX Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VDDPRSEEDKRFESHIECRK 20
XX |||||||:|||||||
XX 915 VDDPRSEEDKRFESHIECRK 934
XX
XX
XX RESULT 14
XX AAE19899
XX ID AAE19899 standard; Protein; 2227 AA.
XX
XX AC AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX
XX Hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX
XX WO200213855-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB01808.
XX
XX 17-AUG-2000; 2000US-225767P.
XX
XX 29-AUG-2000; 2000US-229175P.

PR 03-NOV-2000; 2000US-0705547.
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
XX
XX N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus
XX
XX Claim 11; Page 82-87; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to
XX a hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is hepatitis A virus (HAV) protein.
XX
XX Sequence 2227 AA;
XX
XX Query Match 97.2%; Score 106; DB 23; Length 2227;
XX Best Local Similarity 95.0%; Pred. No. 3e-07;
XX Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VDDPRSEEDKRFESHIECRK 20
XX |||||||:|||||||
XX 779 VDDPRSEEDKRFESHIECRK 798
XX
XX
XX RESULT 15
XX AAP60066
XX ID AAP60066 standard; Protein; 2227 AA.
XX
XX AC AAP60066;
XX
XX 26-JUN-1991 (first entry)
XX
XX Sequence of viral L434 polypeptide encoded by the complete
XX nucleotide sequence of the HAV genome.
XX
XX Diagnosis; vaccine; passive immunotherapy.
XX
XX Hepatitis A virus.
XX
XX Key Location/Qualifiers
XX FH Region 1..245
XX FT /label= P1.1A
XX FT 246..491
XX FT /label= 1B
XX FT 492..836
XX FT /label= 1C
XX FT 837..980
XX FT /label= P2.2A
XX FT 981..1076
XX FT /label= 2B
XX FT 1077..1422
XX FT /label= 2C
XX FT 1423..1484
XX FT /label= P3.3A
XX FT 1485..1507
XX FT /label= 3B
XX FT 1508..1678
XX FT /label= 3C
XX FT 1679..2227
XX FT /label= 3D
XX
XX EP199480-A.

```
XX 29-OCT-1986.
XX
XX 03-APR-1986; 86EP-0302465.
XX
XX 03-APR-1985; 85US-0719329.
XX
XX (CHIR-) CHIRON CORP.
XX
XX Dina D, Potter SJ, Vannest GA, Caput D;
XX WPI; 1986-286213/44.
XX DR N-PSDB; AAN60080.
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use
XX in prodn. of vaccines and diagnostic probes
XX
XX Claim 5; Fig 1; 18pp; English.
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
XX hepatitis A virus; transformed hosts may be used for expression of
XX polypeptides and fragments useful in vaccines without risk of
XX infection by the virus or in prodn. of particles which are capable
XX of inducing immunocompetent B cells for passive immunotherapy. Pref.
XX epitope is derived from AAs 445-657 or 792-848 of the HAV
XX polypeptide sequence (AAP60066).
XX
XX Sequence 2227 AA;
XX
XX Query Match 79.8%; Score 87; DB 7; Length 2227;
XX Best Local Similarity 80.0%; Pred. NO. 0.00024;
XX Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 VDDPRSEEDKRFESHIECRK 20
XX | | | | | | | | | | | | | | | |
XX Db 779 VDDPRSEEDTRFDTHIESRK 798
XX
XX Search completed: February 3, 2003, 13:11:28
XX Job time : 29.8333 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 : Search time 9.5 seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432A-38
Perfect score: 109
Sequence: 1 VDDPRSEEDKRFESHIECRK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTDUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	1091	6	Patent No. 5516630
2	109	100.0	2227	3	Sequence 2, Appli
3	109	100.0	2227	3	Sequence 4, Appli
4	109	100.0	2227	3	Sequence 6, Appli
5	109	100.0	2227	4	Sequence 2, Appli
6	109	100.0	2227	4	Sequence 4, Appli
7	109	100.0	2227	4	Sequence 2, Appli
8	109	100.0	2227	4	Sequence 2, Appli
9	109	100.0	2227	4	Sequence 4, Appli
10	109	100.0	2227	4	Sequence 6, Appli
11	81	74.3	839	1	Sequence 2, Appli
12	44.5	40.8	1334	2	Sequence 2, Appli
13	44.5	40.8	1334	2	Sequence 2, Appli
14	43	39.4	780	2	Patent No. 5273901
15	42	38.5	180	6	Patent No. 5482709
16	42	38.5	180	6	Sequence 5, Appli
17	42	38.5	560	2	Sequence 4664, Ap
18	41	37.6	168	4	Sequence 22, Appl
19	41	37.6	901	3	Sequence 8, Appli
20	41	37.6	906	3	Sequence 10, Appl
21	41	37.6	909	3	Sequence 18, Appl
22	41	37.6	909	3	Sequence 12, Appl
23	41	37.6	909	3	Sequence 14, Appl
24	41	37.6	914	3	Sequence 20, Appl
25	41	37.6	925	4	
26	41	37.6	926	3	
27	41	37.6	926	3	

28	41	37.6	931	3	US-08-936-135-16	Sequence 16, Appl
29	41	37.6	1805	1	US-07-853-913-2	Sequence 2, Appli
30	40	36.7	275	4	US-08-900-565-4	Sequence 4, Appli
31	40	36.7	275	4	US-09-149-534-4	Sequence 4, Appli
32	40	36.7	335	2	US-08-379-556A-10	Sequence 10, Appl
33	40	36.7	407	3	US-08-989-370-5	Sequence 5, Appli
34	39	35.8	146	4	US-08-792-013-7	Sequence 7, Appli
35	39	35.8	147	1	US-08-171-385-5	Sequence 5, Appli
36	39	35.8	147	3	US-08-361-441B-5	Sequence 5, Appli
37	39	35.8	181	4	US-09-134-001C-2892	Sequence 2892, Ap
38	39	35.8	298	4	US-09-424-349A-7	Sequence 7, Appli
39	39	35.8	440	4	US-09-077-955-26	Sequence 26, Appl
40	39	35.8	444	4	US-09-424-283-2	Sequence 2, Appli
41	39	35.8	524	4	US-09-424-283-1	Sequence 1, Appli
42	39	35.8	1001	4	US-09-060-410-2	Sequence 2, Appli
43	39	35.8	1012	4	US-08-811-481-16	Sequence 16, Appl
44	38.5	35.3	544	3	US-08-591-685-9	Sequence 9, Appli
45	38.5	35.3	1349	2	US-08-612-734B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 109; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 1016 VDDPRSEEDKRFESHIECRK 1035

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 7

US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154

COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171.387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-171-387-2

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 8

US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653.499

; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 9

US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653.499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 10

US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653.499
; CURRENT FILING DATE: 2000-08-31

;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||.....
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 11
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 74.3%; Score 81; DB 1; Length 839;
Best Local Similarity 82.4%; Pred. No. 0.00013;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
|||||.....
Db 779 VDDPRFEDRRFESHIE 795

RESULT 12
US-08-996-545-2
; Sequence 2, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-996-545-2

Query Match 40.8%; Score 44.5; DB 2; Length 1334;
Best Local Similarity 55.0%; Pred. No. 95;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 3 DPRSEEDK---RFESHIECR 19
|||.....
Db 413 DPYSNEGKTLDFEGHIELR 432

RESULT 13
US-09-328-320-2
; Sequence 2, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center

; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-328-320-2

Query Match 40.8%; Score 44.5; DB 4; Length 1334;
Best Local Similarity 55.0%; Pred. No. 95;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 3 DPRSEEDK---RFESHIECR 19
||| ||| ||| ||| |||
Db 413 DPYSNEGKTLDFEGHIELR 432

RESULT 14
US-08-887-798-2
; Sequence 2, Application US/08887798
; Patent No. 5922556
; GENERAL INFORMATION:
; APPLICANT: Mayeux, Richard
; APPLICANT: Graziano, Joseph H.
; APPLICANT: Freyer, Greg
; TITLE OF INVENTION: PARKINSON'S DISEASE TESTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,798
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51949/JPW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-887-798-2

Query Match 39.4%; Score 43; DB 2; Length 780;
Best Local Similarity 42.1%; Pred. No. 90;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECR 19
: ||| | : | : | : | : |
Db 74 LDDPASQEIERGKSYLRRLR 92

RESULT 15
5273901-7
; Patent No. 5273901
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS
; SPOOROZOITE 21.5 KB ANTIGEN, AC-6B
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/581,693
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 215,162
; FILING DATE: 05-JUL-1988
; APPLICATION NUMBER: 746,520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627,811
; FILING DATE: 05-JUL-1984
; SEQ ID NO: 7:
; LENGTH: 180
5273901-7

Query Match 38.5%; Score 42; DB 6; Length 180;
Best Local Similarity 56.2%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 DDPRESEEDKRFESHIE 17
: ||| ||| ||| | | |
Db 3 EDKREEDKREEEEEEE 18

Search completed: February 3, 2003, 13:16:36
Job time : 11.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:14:27 : Search time 6 Seconds
(without alignments)
67.262 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDPRSEEDKRFESHIECRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	2227	9	US-10-135-988-2
2	109	100.0	2227	9	US-10-135-988-4
3	109	100.0	2227	9	US-10-135-988-6
4	106	97.2	2227	9	US-10-104-966-12
5	106	97.2	2227	10	US-09-929-955-12
6	44.5	40.8	1334	9	US-09-758-828-2
7	41	37.6	931	12	US-10-104-440-4
8	39.5	36.2	1289	9	US-09-932-145-11
9	39	35.8	60	10	US-09-864-761-46320
10	39	35.8	146	10	US-09-782-142-7
11	39	35.8	172	9	US-09-922-199A-26
12	39	35.8	440	9	US-10-016-283-26
13	39	35.8	1012	10	US-09-876-527-16
14	39	35.8	2150	10	US-09-321-987B-2
15	39	35.8	2165	10	US-09-800-729-155
16	38	34.9	142	9	US-09-764-868-1126
17	38	34.9	145	10	US-09-923-995-2
18	38	34.9	164	9	US-09-738-626-6141
19	38	34.9	219	9	US-09-764-868-715

20	38	34.9	282	10	US-09-864-761-38121	Sequence 38121, A
21	38	34.9	659	9	US-10-144-621-2	Sequence 2, Appli
22	38	34.9	1924	9	US-09-866-557A-2	Sequence 2, Appli
23	38	34.9	2485	10	US-09-802-669-46	Sequence 46, Appl
24	37.5	34.4	1164	10	US-09-950-046A-2	Sequence 2, Appli
25	37	33.9	74	10	US-09-864-761-38928	Sequence 38928, A
26	37	33.9	76	9	US-09-749-637A-68	Sequence 68, Appl
27	37	33.9	81	10	US-09-864-761-40358	Sequence 40358, A
28	37	33.9	88	9	US-09-965-528-2	Sequence 2, Appli
29	37	33.9	150	10	US-09-864-761-39141	Sequence 39141, A
30	37	33.9	223	9	US-10-028-072-30	Sequence 30, Appl
31	37	33.9	308	9	US-10-028-072-100	Sequence 100, App
32	37	33.9	342	10	US-09-925-300-1343	Sequence 1343, Ap
33	37	33.9	354	10	US-09-815-242-5345	Sequence 5345, Ap
34	37	33.9	394	10	US-09-815-242-5574	Sequence 5574, Ap
35	37	33.9	394	10	US-09-815-242-12442	Sequence 12442, A
36	37	33.9	423	10	US-09-815-242-5063	Sequence 5063, Ap
37	37	33.9	444	10	US-09-815-242-12483	Sequence 12483, A
38	37	33.9	487	10	US-09-745-763-108	Sequence 108, App
39	37	33.9	599	10	US-09-952-013A-4	Sequence 4, Appli
40	37	33.9	609	10	US-09-864-761-33429	Sequence 33429, A
41	37	33.9	761	9	US-10-045-792-13	Sequence 13, Appl
42	37	33.9	970	10	US-09-901-419-2	Sequence 2, Appli
43	37	33.9	983	12	US-10-047-757-2	Sequence 2, Appli
44	36.5	33.5	555	10	US-09-764-864-1161	Sequence 1161, Ap
45	36.5	33.5	555	10	US-09-764-864-1574	Sequence 1574, Ap

ALIGNMENTS

RESULT 1
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 109; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
DB 779 VDDPRSEEDKRFESHIECRK 798

RESULT 2
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H

```
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match      100.0%; Score 109; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
   |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 3
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match      100.0%; Score 109; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
   |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/7705,547
```

```
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      97.2%; Score 106; DB 9; Length 2227;
Best Local Similarity 95.0%; Pred. No. 4.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
   |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/7705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      97.2%; Score 106; DB 10; Length 2227;
Best Local Similarity 95.0%; Pred. No. 4.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
   |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 6
US-09-758-828-2
; Sequence 2, Application US/09758828
; Patent No. US20020164772A1
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
```

STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/758,828
FILING DATE: 11-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1334 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-758-828-2

Query Match 40.8%; Score 44.5; DB 9; Length 1334;
Best Local Similarity 55.0%; Pred. No. 66;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 3 DPRSEEDK---RFESHIECR 19
||| | | | | | | | | | |
Db 413 DPYSEGTLDHFEGHIELR 432

RESULT 7
US-10-104-440-4
; Sequence 4, Application US/10104440
; Patent No. US20020132774A1
; GENERAL INFORMATION:
; APPLICANT: KLAGSBRUN, Michael
; APPLICANT: SOKER, Shay
; APPLICANT: MIAO, Hua Quan
; TITLE OF INVENTION: ANTAGONISTS OF NEUROPILIN RECEPTOR FUNCTION AND USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 48802 C
; CURRENT APPLICATION NUMBER: US/10/104,440
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/580,803
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 60/069,155
; PRIOR FILING DATE: 1997-12-09
; PRIOR APPLICATION NUMBER: 60/069,687
; PRIOR FILING DATE: 1997-12-29
; PRIOR APPLICATION NUMBER: 60/078,541
; PRIOR FILING DATE: 1998-03-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: human
US-10-104-440-4
Query Match 37.6%; Score 41; DB 12; Length 931;
Best Local Similarity 37.5%; Pred. No. 1.5e+02;

Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
Qy 1 VDDPRSEEDKRFESHI 16
: ||||::| | |::
Db 538 IODPRTQPKLFEQNM 553
RESULT 8
US-09-932-145-11
; Sequence 11, Application US/09932145
; Patent No. US20020161191A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Mintier, Gabe
; APPLICANT: Kinney, Gene G
; APPLICANT: Ramanathan, Chandra S
; TITLE OF INVENTION: NOVEL IMIDAZOLINE RECEPTOR HOMOLOGS
; FILE REFERENCE: D0020 NP
; CURRENT APPLICATION NUMBER: US/09/932,145
; CURRENT FILING DATE: 2001-08-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1289
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: DATABASE ACCESSION NUMBER: Database: Genbank,
; OTHER INFORMATION: Accession: AAF52305
US-09-932-145-11

Query Match 36.2%; Score 39.5; DB 9; Length 1289;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 3 DPRSE-EDKRFESHIECRK 20
: |||| |:: | | | |:
Db 430 NPRSEYEEDPNSHLETKK 448

RESULT 9
US-09-864-761-46320
; Sequence 46320, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 46320
;; LENGTH: 60
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC011375.3
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
;; OTHER INFORMATION: SWISSPROT HIT: Q53213, EVALUE 3.10e-01
;; OTHER INFORMATION: EST_HUMAN HIT: AI218274.1, EVALUE 1.00e-28
US-09-864-761-46320

Query Match 35.8%; Score 39; DB 10; Length 60;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 RSEEDKRFESHIE 17
| | | | | | | | | |
DB 29 REEKEKLFNEHIE 41

RESULT 10
US-09-782-142-7
; Sequence 7, Application US/09782142
; Patent No. US20020001827A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; ; Hawkins, Phillip R.
; ; Murry, Lynn E.
; ; Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CYTOKINES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/782,142
; FILING DATE: 12-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/792,013
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0205 US

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 146 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 1514969
;; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-782-142-7

Query Match 35.8%; Score 39; DB 10; Length 146;
Best Local Similarity 35.0%; Pred. No. 39;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
: | | | : | : | : | : |
DB 31 LDDPKYSSDEDLQSKLEAFK 50

RESULT 11
US-09-922-199A-26
; Sequence 26, Application US/09922199A
; Publication No. US20020187138A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 15368, A NOVEL HUMAN GTP-RELEASING
; FILE OF INVENTION: FACTOR FAMILY MEMBER AND USES THEREFOR
; FILE REFERENCE: 38152002500
; CURRENT APPLICATION NUMBER: US/09/922,199A
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/222,622
; PRIOR FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-922-199A-26

Query Match 35.8%; Score 39; DB 9; Length 172;
Best Local Similarity 53.3%; Pred. No. 47;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 SEEDKRFESHIECRK 20
| | | | | | | | | |
DB 112 SEEKERSEDLLECKTK 126

RESULT 12
US-10-016-283-26
; Sequence 26, Application US/10016283
; Patent No. US20020164702A1
; GENERAL INFORMATION:
; APPLICANT: Valenzuela et al., David M.
; TITLE OF INVENTION: NOVEL TYROSINE KINASE RECEPTORS AND LIGANDS
; FILE REFERENCE: REG195-B-PCT-US
; CURRENT APPLICATION NUMBER: US/10/016,283
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US/09/077,955A
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: PCT/US96/20696
; PRIOR FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26

LENGTH: 440
TYPE: PRT
ORGANISM: Homo sapiens
US-10-016-283-26

Query Match 35.8%; Score 39; DB 9; Length 440;
Best Local Similarity 42.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECR 19
| | : : : | | | : |
Db 14 VADYKDDDDKNGFSLHLR 32

RESULT 13
US-09-876-527-16
Sequence 16, Application US/09876527
Patent No. US20020102616A1
GENERAL INFORMATION:
APPLICANT: Kindsvogel, Wayne
Jelinek, Laura J.
Sheppard, Paul O.
Hagopian, William A.
LaGasse, James M.

TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/876,527
FILING DATE: 07-Jun-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/811,481
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Lingenfelter, Susan
REGISTRATION NUMBER: P-41,156
REFERENCE/DOCKET NUMBER: 95-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6675
TELEFAX: 206-442-6678
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Query Match 35.8%; Score 39; DB 10; Length 1012;
Best Local Similarity 53.3%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPRESEEDKRFESHI 16
| | : : : | | | : |
Db 175 DRPRAEGDDRFKSKI 189

RESULT 14
US-09-321-987B-2
Sequence 2, Application US/09321987B
Patent No. US20020102210A1
GENERAL INFORMATION:
APPLICANT: Kimble, Judith E
APPLICANT: Billeloch, Robert H
TITLE OF INVENTION: Agent and Method for Modulating Cell Migration
FILE REFERENCE: 960296.95386
CURRENT APPLICATION NUMBER: US/09/321,987B
CURRENT FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,170
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/129,023
PRIOR FILING DATE: 1999-04-13
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2150
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-321-987B-2

Query Match 35.8%; Score 39; DB 10; Length 2150;
Best Local Similarity 33.3%; Pred. No. 7.8e+02;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 5 RSEEDKRFESHIECR 19
| | : : : | | | : |
Db 505 RDEPGKKYDAHQCK 519

RESULT 15
US-09-800-729-155
Sequence 155, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 155
LENGTH: 2165
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-155

Query Match 35.8%; Score 39; DB 10; Length 2165;
Best Local Similarity 33.3%; Pred. No. 7.8e+02;
Matches 5; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 5 RSEEDKRFESHIECR 19
| | : : : | | | : |
Db 520 RDEPGKKYDAHQCK 534

Search completed: February 3, 2003, 13:32:24
Job time : 7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 : Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDPRSEEDKRFESHIECRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	100.0	56	2 PQ0431	genome polyprotein
2	109	100.0	319	2 JH0135	genome polyprotein
3	109	100.0	1358	2 A03905	genome polyprotein
4	109	100.0	2227	1 GNNYHM	genome polyprotein
5	109	100.0	2227	1 GNNYMK	genome polyprotein
6	109	100.0	2227	1 GNNYHB	genome polyprotein
7	106	97.2	56	2 PQ0427	genome polyprotein
8	106	97.2	56	2 PQ0429	genome polyprotein
9	106	97.2	341	2 S04137	genome polyprotein
10	106	97.2	852	1 GNNYHA	genome polyprotein
11	106	97.2	2227	1 GNNYHR	genome polyprotein
12	99	90.8	56	2 PQ0430	genome polyprotein
13	96	88.1	56	2 PQ0432	genome polyprotein
14	95	87.2	56	2 PQ0428	genome polyprotein
15	91	83.5	56	2 PQ0434	genome polyprotein
16	84	77.1	56	2 PQ0437	genome polyprotein
17	84	77.1	56	2 PQ0436	genome polyprotein
18	84	77.1	2230	1 GNNYSA	genome polyprotein
19	81	74.3	55	2 PQ0433	genome polyprotein
20	81	74.3	55	2 PQ0435	genome polyprotein
21	81	74.3	839	1 GNNYS2	genome polyprotein
22	60	55.0	300	2 A3327	genome polyprotein
23	51	46.8	1548	2 T25808	genome polyprotein
24	50.5	46.3	754	2 T25551	hypothetical prote
25	50	45.9	1663	1 C3RT	complement C3 prec
26	50	45.9	1666	1 C3GP	complement C3 prec
27	48	44.0	322	2 A32329	complement C3 - Af
28	48	44.0	564	2 S37241	legumin B - fava b
29	46.5	42.7	454	2 B84699	hypothetical prote

30 46 42.2 242 2 E83077 conserved hypothet
31 46 42.2 308 2 T17524 hypothetical prote
32 46 42.2 1663 1 C3MS complement C3 prec
33 45.5 41.7 650 2 T36419 hypothetical prote
34 45 41.3 209 2 T20155 hypothetical prote
35 45 41.3 480 2 T20154 hypothetical prote
36 45 41.3 586 1 E69314 replication licens
37 45 41.3 741 2 S73827 hypothetical prote
38 44 40.4 222 2 E71507 probable L4 riboso
39 44 40.4 234 2 B83065 conserved hypothet
40 44 40.4 433 2 A23850 phosphopyruvate hy
41 44 40.4 490 2 C84091 hypothetical prote
42 44 40.4 499 2 T04730 cytochrome P450 ho
43 44 40.4 593 2 A81021 para-aminobenzoate
44 44 40.4 598 2 D81965 probable para-amin
45 44 40.4 782 2 G96698 hypothetical prote

ALIGNMENTS

RESULT 1

PQ0431

genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)

C;Species: human hepatitis A virus

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C;Accession: PQ0431

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wl

J. gen. Virol. 73, 1365-1377, 1992

A;Title: Genetic relatedness of hepatitis A virus strains recovered from different ge

A;Reference number: PQ0427; MUID:92300330; PMID:1318940

A;Accession: PQ0431

A;Molecule type: mRNA

A;Residues: 1-56 <ROB>

A;Note: this protein is from the VP1/2A Junction region

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 109; DB 2; Length 56;

Best Local Similarity 100.0%; Pred. No. 4.9e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20

|||||

Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 2

JH0135

genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)

N;Contains: amino end of core protein 2A; coat protein 1D

C;Species: human hepatitis A virus

A;Note: host Homo sapiens (man)

C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994

C;Accession: JH0135

R;Robertson, B.H.; Brown, V.K.; Bradley, D.W.

Virus Res. 8, 305-316, 1987

A;Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus

A;Reference number: JH0135; MUID:88129044; PMID:2829458

A;Accession: JH0135

A;Molecule type: genomic RNA

A;Residues: 1-319 <ROB>

A;Cross-references: GB:M22821

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; glycoprotein; polyprotein

F;1-300/Product: coat protein 1D #status predicted <CPD>

F;301-319/Product: core protein 2A (fragment) #status predicted <C2A>

F;237/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 109; DB 2; Length 319;

Best Local Similarity 100.0%; Pred. No. 3.1e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 288 VDDPRSEEDKRFESHIECRK 307

RESULT 3
A03905
genome polypotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein 2A; protein 2B; protein 2C
C:Species: human hepatitis A virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C:Accession: A03905
R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstone, Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:Reference number: A03905; MUID:85166289; PMID:2984684
A:Accession: A03905
A:Molecule type: genomic RNA
A:Residues: 1-1358 <BAR>
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; polypotein
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>
F:855-1358/Product: protein 2B (RNA polymerase) (fragment) #status predicted <C2B>
F:855-1358/Product: protein 2C (RNA polymerase) (fragment) #status predicted <C2C>
F:855-1358/Product: protein 2D (RNA polymerase) (fragment) #status predicted <C2D>

Query Match 100.0%; Score 109; DB 1; Length 1358;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4
GNNYHM
genome polypotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M. J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with other HAV sequences
Reference number: A25981; MUID:87061253; PMID:3023706
Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 5
GNNYMK
genome polypotein - human hepatitis A virus (strain HM-175/WMK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with other HAV sequences
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 6
GNNYHB
genome polypotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Kiehn, R.; Wimmer, E.; Deinhard Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso 179)
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1108/Product: core protein 2B #status predicted <C2B>
F:1109-1438/Product: core protein 2C #status predicted <C2C>
F:1439-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1520-1736/Product: genome-linked protein VPg #status predicted <VPg>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 7

PQ0427
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0427; MUID:92300330; PMID:1318940

A:Accession: PQ0427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Comment: This protein is from the VP1/2A Junction region.
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 97.2%; Score 106; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.4e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 8

PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0427; MUID:92300330; PMID:1318940

A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 97.2%; Score 106; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.4e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 9

S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Antonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A:Reference number: S04137; MUID:89263805; PMID:2542903

A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:962301; PIDN:CAA32794.1; PID:94377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 97.2%; Score 106; DB 2; Length 341;
Best Local Similarity 95.0%; Pred. No. 9.7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 283 VDDPRSEEDKRFESHIECRK 302

RESULT 10

GNVYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03904
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793

A:Accession: A03904
A:Molecule type: genomic RNA
A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <CIA>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 97.2%; Score 106; DB 1; Length 852;
Best Local Similarity 95.0%; Pred. No. 2.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 11

GNVYHR
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; MUID:85190549; PMID:29866127

A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>

A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr
F:1-245/Product: coat protein 1A #status predicted <CIA>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-980/Product: core protein 2A #status predicted <C2A>
F:1076-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 97.2%; Score 106; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 7.1e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 12

P00430
genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VPI/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 90.8%; Score 99; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 1.7e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

RESULT 13

P00432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00432
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VPI/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 88.1%; Score 96; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 5.1e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

RESULT 14

P00428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00428
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VPI/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 87.2%; Score 95; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 7.3e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

RESULT 15

P00434
genome polyprotein - human hepatitis A virus (strain KPH) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00434
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00434
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VPI/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 83.5%; Score 91; DB 2; Length 56;
Best Local Similarity 85.0%; Pred. No. 3e-07;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

Search completed: February 3, 2003, 13:15:32
Job time : 11.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:05:12 ; Search time 5.16667 Seconds
(without alignments)
160.554 Million cell updates/sec

Title: US-09-171-432a-38

Perfect score: 109

Sequence: 1 VDDPRSEEDKRFESHIKCRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	100.0	2226	1	POLG_HPAV2
2	109	100.0	2226	1	POLG_HPAV4
3	109	100.0	2226	1	POLG_HPAV8
4	109	100.0	2227	1	POLG_HPAVH
5	109	100.0	2227	1	POLG_HPAVM
6	106	97.2	341	1	POLG_HPAV1
7	106	97.2	852	1	POLG_HPAVC
8	106	97.2	2227	1	POLG_HPAVL
9	88	80.7	808	1	POLG_HPAVG
10	84	77.1	2230	1	POLG_HPAVS
11	81	74.3	839	1	POLG_HPAVT
12	50	45.9	1663	1	CO3_RAT
13	50	45.9	1666	1	CO3_CAVPO
14	48	44.0	323	1	CO3_XENIA
15	46	42.2	242	1	Y9E3_PSEAE
16	46	42.2	1663	1	CO3_MOUSE
17	45	41.3	741	1	YD35_MYCPN
18	44	40.4	222	1	RL4_CHLTR
19	44	40.4	1101	1	FURC_DRONE
20	43.5	39.9	801	1	DHGA_ACICA
21	43	39.4	265	1	GCTB_ACIFE
22	43	39.4	478	1	RPB1_EUPOC
23	43	39.4	502	1	C911_ARATH
24	43	39.4	780	1	ACON_HUMAN
25	43	39.4	963	1	YQ36_CAEEL
26	43	39.4	2137	1	SPCB_HUMAN
27	42	38.5	269	1	SFAS_CHLEU
28	42	38.5	276	1	SFAS_CHLEU
29	42	38.5	408	1	RTCB_ECOLI
30	42	38.5	464	1	AK15_RAT
31	42	38.5	559	1	CO9_HUMAN
32	42	38.5	679	1	YIS3_YEAST
33	42	38.5	918	1	IF38_SCHPO

34	41	37.6	267	1	SURE_METAC
35	41	37.6	396	1	EFTU_FLESI
36	41	37.6	434	1	YK12_YEAST
37	41	37.6	492	1	PROD_SCHPO
38	41	37.6	571	1	YB63_SCHPO
39	41	37.6	691	1	Y104_YEAST
40	41	37.6	925	1	NRP2_RAT
41	41	37.6	931	1	NRP2_HUMAN
42	41	37.6	931	1	NRP2_MOUSE
43	41	37.6	1021	1	TSCC_HUMAN
44	41	37.6	1805	1	NEST_RAT
45	41	37.6	1882	1	POL2_TRSVR

ALIGNMENTS

RESULT 1
POLG_HPAV2
ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 24a).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91162758; PubMed-1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination".
RT J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59810; AAA45468.1; --
CC MEROPS: C03.005;
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR001205; RNA_pol_P3D.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 25 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.

FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 109; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 2
POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID POLG_HPAV4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 43c).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -I- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59809; AAA45469.1; -
CC MEROPS; C03.005; -
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Polyprotein; Coat protein; Core protein; Thiol protease;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).
FT CHAIN 795 900 COAT PROTEIN VP1 (PID).
FT CHAIN 901 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 COAT PROTEIN VP2 (PIB).
FT CHAIN 1496 1518 COAT PROTEIN VP3 (PIC).
FT CHAIN 1519 1737 PROBABLE PROTEIN P3A.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3B.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BF75 CRC64;

Query Match 100.0%; Score 109; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 3
POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID POLG_HPAV8
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 18f).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -I- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59808; AAA45467.1; -
CC MEROPS; C03.005; -
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR001205; RNA_pol_P3D.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC Polyprotein; Coat protein; Core protein; Thiol protease;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).
FT CHAIN 795 900 COAT PROTEIN VP1 (PID).
FT CHAIN 901 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 COAT PROTEIN VP2 (PIB).
FT CHAIN 1496 1518 COAT PROTEIN VP3 (PIC).
FT CHAIN 1519 1737 PROBABLE PROTEIN P3A.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3B.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 109; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 VDDPRSEEDKRFESHIECRK 20
    |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4
POLG_HPAAVH STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Attenuated;
RX MEDLINE=8715701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstein S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: STRAIN HM-175/77 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M14114; AAA45475.1; -.
CC EMBL; M14707; AAA45465.1; -.
CC EMBL; M14707; AAA45466.1; ALT_INIT.
CC EMBL; M16632; AAA45471.1; -.
CC PIR; A25981; GNNYHM.
CC PIR; A25914; GNNYMK.
CC PIR; A03905; A03905.

```

```

DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
    Query Match 100.0%; Score 109; DB 1; Length 2227;
    Best Local Similarity 100.0%; Pred. No. 1.5e-08;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 VDDPRSEEDKRFESHIECRK 20
    |||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 5
POLG_HPAAVH STANDARD; PRT; 2227 AA.
AC P13901; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
RA Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MBB).";
RL Virus Res. 8:153-171(1987).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

```



```
RESULT 8
POLG_HP AVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)]
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dina D.;
RA "Primary structure and gene organization of human hepatitis A virus.";
RA Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: K02990; AAA45472.1;
DR PIR: A03903; GNHYR.
DR MEROPS: C03.005;
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
FT CHAIN 837 980 CORE PROTEIN P2A.
FT CHAIN 981 1076 CORE PROTEIN P2B.
FT CHAIN 1077 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1484 PROBABLE PROTEIN P3A.
FT CHAIN 1485 1507 PROBABLE PROTEIN P3B.
FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 97.2%; Score 106; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 4.4e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
| | | | | | | | | | | | | | | | | |
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 9
POLG_HP AVL STANDARD; PRT; 808 AA.
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
```

```
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain GAY6).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RA "Characterization of a genetic variant of human hepatitis A virus.";
RA J. Med. Virol. 36:118-124(1992).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M65695; AAA45477.1;
DR Polyprotein; Coat protein; Core protein.
FT CHAIN 1 1
FT CHAIN <1 2 COAT PROTEIN VP4 (PIA).
FT CHAIN 3 223 COAT PROTEIN VP2 (PIB).
FT CHAIN 224 470 COAT PROTEIN VP3 (PIC).
FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
FT CHAIN 771 >808 CORE PROTEIN P2A.
FT NON_TER 808 808
SQ SEQUENCE 808 AA; 908332 MW; D80CE7E57A479C12 CRC64;

Query Match 80.7%; Score 88; DB 1; Length 808;
Best Local Similarity 80.0%; Pred. No. 8.3e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
| | | | | | | | | | | | | | | | | |
Db 758 VDDPRTEEDKRFESHIECRK 777

RESULT 10
POLG_HP AVL STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)]
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RA structure and growth in cell culture with other HAV strains.";
RA J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
```


CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC BASOPHILIC LEUKOCYTES.
CC
CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
CC RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,
CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA
CC CHAIN).
CC
CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC
CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X52477; CAA36716.1; -
CC EMBL: M29866; AAA40837.1; ALT_SEQ.
CC PIR: A01260; A01260.
CC PIR: S15764; S15764.
CC HSSP: P01024; 1C3D.
CC InterPro: IPR002890; A2M_N.
CC InterPro: IPR000020; Anaphylatoxin.
CC InterPro: IPR001840; Anaphylatoxn.
CC InterPro: IPR001599; MacrogloblnA2.
CC InterPro: IPR001134; Netrin_C.
CC Pfam: PF00207; A2M; 1.
CC Pfam: PF01759; NTR; 1.
CC Pfam: PF01821; ANATO; 1.
CC Pfam: PF01835; A2M_N; 1.
CC PRINTS: PR00004; ANAPHYLATOXN.
CC ProDom: PD003284; Anaphylatoxin; 1.
CC SMART: SM00104; ANATO; 1.
CC PROSITE: PS00477; ALPHA_2_MACROGLOBULIN; 1.
CC PROSITE: PS01177; ANAPHYLATOXIN.1; 1.
CC PROSITE: PS01178; ANAPHYLATOXIN.2; 1.
CC Complement pathway; Complement alternate pathway; Plasma;
CC Inflammatory response; Glycoprotein; Signal.
CC
CC SIGNAL 1 24
CC CHAIN 25 1663 COMPLEMENT C3.
CC CHAIN 25 666 BETA CHAIN.
CC CHAIN 671 1663 ALPHA CHAIN.
CC PEPTIDE 671 748 C3A ANAPHYLATOXIN.
CC CHAIN 749 1663 C3B (ALPHA' CHAIN).
CC SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).
CC DOMAIN 693 728 ANAPHYLATOXIN-LIKE.
CC DISULFID 558 816 INTERCHAIN (BY SIMILARITY).
CC DISULFID 626 661 BY SIMILARITY.
CC DISULFID 693 720 BY SIMILARITY.
CC DISULFID 694 727 BY SIMILARITY.
CC DISULFID 707 728 BY SIMILARITY.
CC DISULFID 873 1513 BY SIMILARITY.
CC DISULFID 1101 1158 BY SIMILARITY.
CC DISULFID 1358 1489 BY SIMILARITY.
CC DISULFID 1389 1458 BY SIMILARITY.
CC DISULFID 1506 1511 BY SIMILARITY.
CC DISULFID 1518 1590 BY SIMILARITY.
CC DISULFID 1537 1661 BY SIMILARITY.
CC THIOLEST 1010 1013
CC CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (PROBABLE).
CC CARBOHYD 1617 1617 N-LINKED (GLCNAC. . .) (PROBABLE).
CC CONFLICT 721 722 LK -> KL (IN REF. 2).
CC SEQUENCE 1663 AA; 186460 MW; 2F87CCB143CDD4BC CRC64;

Query Match 45.9%; Score 50; DB 1; Length 1663;
Best Local Similarity 38.9%; Pred. No. 11;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIECR 19
I: : : : :
Db 1574 DEVQAGQERRFISHVKCR 1591

RESULT 13
CO3_CAVPO
ID CO3_CAVPO STANDARD; PRT; 1666 AA.
AC P12387;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Complement C3 precursor [Contains: C3A anaphylatoxin].
GN C3.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90307998; PubMed=1973176;
RA Auerbach H.S., Burger R., Dodds A., Colten H.R.;
RT "Molecular basis of complement C3 deficiency in guinea pigs.";
RL J. Clin. Invest. 86:96-106(1990).
RN [2]
RP SEQUENCE OF 676-753.
RX MEDLINE=89113342; PubMed=3064079;
RA Gerard N.P., Lively M.O., Gerard C.;
RT "Amino acid sequence of guinea pig C3a anaphylatoxin.";
RL Protein Seq. Data Anal. 1:473-478(1988).
RN [3]
RP SEQUENCE OF 993-1032.
RX MEDLINE=83178889; PubMed=6838833;
RA Thomas M.L., Tack B.F.;
RT "Identification and alignment of a thiol ester site in the third
RP component of guinea pig complement.";
RL Biochemistry 22:942-947(1983).
CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC BASOPHILIC LEUKOCYTES.
CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
CC RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,
CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA
CC CHAIN).
CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC -1- SIMILARITY: CONTAINS 1 ANAPHYLATOXIN-LIKE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: M34054; AAA37038.1; -
CC PIR: A37156; A37156.
CC PIR: S03375; S03375.
CC PIR: D20342; D20342.
CC HSSP: P01024; 1C3D.
CC InterPro: IPR002890; A2M_N.
CC InterPro: IPR000020; Anaphylatoxin.
CC InterPro: IPR001840; Anaphylatoxn.
CC InterPro: IPR001599; MacrogloblnA2.
CC InterPro: IPR001134; Netrin_C.

RT "PILS and PILR, a two-component transcriptional regulatory system
 RT controlling expression of type 4 fimbriae in Pseudomonas
 RL aeruginosa";
 CC Mol. Microbiol. 7:669-682(1993).
 CC -!- SIMILARITY: BELONGS TO THE UPF0124 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE004868; AAG07931.1; -;
 DR EMBL; L06013; AAA87639.1; -;
 DR InterPro: IPR003730; DUF152.
 DR Pfam; PF02578; DUF152; 1.
 DR TIGRFAMs; TIGR00726; DUF152; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 242 AA; 26047 MW; AD4470FF676FF51D CRC64;

Query Match 42.2%; Score 46; DB 1; Length 242;
 Best Local Similarity 45.0%; Pred. No. 5.9;
 Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 DDPRSEED--KRFESHIECR 19
 ||||:|:|:
 Db 42 DDPRAVEENRRRLTERLECR 61

Search completed: February 3, 2003, 13:12:07
 Job time : 7.16667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:07:47 ; Search time 21.5 Seconds
(without alignments)
191.672 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDPRSEEDKRFSHIECRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	109	100.0	56	12	039865 hepatitis a
2	109	100.0	56	12	039866 hepatitis a
3	109	100.0	56	12	039867 hepatitis a
4	109	100.0	56	12	039868 hepatitis a
5	109	100.0	56	12	039869 hepatitis a
6	109	100.0	56	12	039873 hepatitis a
7	109	100.0	56	12	039874 hepatitis a
8	109	100.0	56	12	039875 hepatitis a
9	109	100.0	56	12	039876 hepatitis a
10	109	100.0	56	12	039877 hepatitis a
11	109	100.0	56	12	039878 hepatitis a
12	109	100.0	56	12	039879 hepatitis a
13	109	100.0	56	12	039880 hepatitis a
14	109	100.0	56	12	039881 hepatitis a
15	109	100.0	56	12	039882 hepatitis a
16	109	100.0	56	12	039883 hepatitis a

17	109	100.0	56	12	067820	hepatitis a
18	109	100.0	56	12	067821	hepatitis a
19	109	100.0	56	12	067822	hepatitis a
20	109	100.0	56	12	067823	hepatitis a
21	109	100.0	56	12	09QCT2	hepatitis a
22	109	100.0	97	12	09QCT1	hepatitis a
23	109	100.0	97	12	09QCT0	hepatitis a
24	109	100.0	116	12	092941	hepatitis a
25	109	100.0	132	12	08V4N6	hepatitis a
26	109	100.0	132	12	08V4N3	hepatitis a
27	109	100.0	132	12	08V4N0	hepatitis a
28	109	100.0	132	12	08V4M2	hepatitis a
29	109	100.0	132	12	08V4M4	hepatitis a
30	109	100.0	132	12	08V4M1	hepatitis a
31	109	100.0	132	12	08V4L9	hepatitis a
32	109	100.0	132	12	08V4L6	hepatitis a
33	109	100.0	132	12	08V4K8	hepatitis a
34	109	100.0	132	12	08V4K5	hepatitis a
35	109	100.0	132	12	08V4K3	hepatitis a
36	109	100.0	132	12	08V4K0	hepatitis a
37	109	100.0	132	12	08V4J9	hepatitis a
38	109	100.0	132	12	08V4J5	hepatitis a
39	109	100.0	258	12	Q99IS7	hepatitis a
40	109	100.0	819	12	Q8V4I2	hepatitis a
41	109	100.0	819	12	Q8V4I1	hepatitis a
42	109	100.0	819	12	Q8V4I0	hepatitis a
43	109	100.0	819	12	Q8V4O9	hepatitis a
44	109	100.0	819	12	Q8V4O8	hepatitis a
45	109	100.0	819	12	Q8V4O7	hepatitis a

ALIGNMENTS

RESULT 1

O39865 PRELIMINARY; PRT; 56 AA.
AC O39865;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus
OX NCBI_taxid=i2092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=412991;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RL 1982-1996";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68689; AAB53585.1; .
DR InterPro; IPR000886; ERTarget.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 4.3e-10; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VDDPRSEEDKRFSHIECRK 20

|||||

Db 16 VDDPRSEEDKRFSHIECRK 35

RESULT 2

O39866 PRELIMINARY; PRT; 56 AA.
ID O39866

```
AC O39866;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96001190;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68690; AAB53586.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 3
O39867
ID O39867 PRELIMINARY; PRT; 56 AA.
AC O39867;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JVR;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68692; AAB53588.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 4
O39868
ID O39868 PRELIMINARY; PRT; 56 AA.
AC O39868;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96001190;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68692; AAB53588.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 5
O39869
ID O39869 PRELIMINARY; PRT; 56 AA.
AC O39869;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96002382;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68694; AAB53590.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 6
O39873
ID O39873 PRELIMINARY; PRT; 56 AA.
AC O39873;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
```

```
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=503712;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68693; AAB53589.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 5
O39869
ID O39869 PRELIMINARY; PRT; 56 AA.
AC O39869;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96002382;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68694; AAB53590.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 6
O39873
ID O39873 PRELIMINARY; PRT; 56 AA.
AC O39873;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
```



```
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2333;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL: U68698; AAB53594.1; -.
DR InterPro: IPR000886; ER_target.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 56
FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 7
Q39874 ID Q39874 PRELIMINARY; PRT; 56 AA.
AC Q39874;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=314274;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
  1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL: U68699; AAB53595.1; -.
DR InterPro: IPR000886; ER_target.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 56
FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 8
Q91PAB ID Q91PAB PRELIMINARY; PRT; 56 AA.
AC Q91PAB;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A40;
RA Fujiwara K.;
RT "hepatitis A virus VP1/2A junction.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046902; BAB08059.1; -.
DR InterPro: IPR000886; ER_target.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 56
FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 9
Q91P95 ID Q91P95 PRELIMINARY; PRT; 56 AA.
AC Q91P95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A7;
RA Fujiwara K.;
RT "hepatitis A virus VP1/2A junction.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046915; BAB08072.1; -.
DR InterPro: IPR000886; ER_target.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1 56
FT SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 10
Q918Q4 ID Q918Q4 PRELIMINARY; PRT; 56 AA.
AC Q918Q4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
```

```
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RJ-004;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil.";
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410380; AAL10176.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 11
Q918Q3 PRELIMINARY; PRT; 56 AA.
AC Q918Q3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RJ-005;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil.";
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410381; AAL10177.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 12
Q918Q1 PRELIMINARY; PRT; 56 AA.
AC Q918Q1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RJ-504;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
```

```
RC STRAIN-RJ-NSG2;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil.";
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410383; AAL10179.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 698E074846AEF4B7 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 13
Q918Q0 PRELIMINARY; PRT; 56 AA.
AC Q918Q0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RJ-055;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil.";
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410384; AAL10180.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6598 MW; E838C51846AEF4B4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 14
Q918P9 PRELIMINARY; PRT; 56 AA.
AC Q918P9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RJ-504;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
```

```
RT and IB in Rio de Janeiro, Brazil.";
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410385; AAL10181.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6598 MW; E838C51846AEF4B4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 15
Q67818 PRELIMINARY; PRT; 56 AA.
D 067818
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-05;
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
RL Israel, based on their VP1/2A nucleotide sequence.";
DR EMBL; Z77243; CAB01036.1; -.
DR InterPro; IPR000886; ER_target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

Search completed: February 3, 2003, 13:14:22
Job time : 22.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 ; Search time 27.8333 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101

Sequence: 1 VLPFPRKMKGLFSQAKISLF 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	AAW42925	Immunogenic Hepati
2	101	100.0	21	AAW42925	Synthetic HAV P2A
3	101	100.0	366	AAP50230	Sequence of hepati
4	101	100.0	854	AAP50287	Sequence encoded b
5	101	100.0	993	AAP50116	Sequence of Hepati
6	101	100.0	993	AAP50231	Sequence encoded b
7	101	100.0	1077	AAW95559	A partial hepatiti
8	101	100.0	1091	AAW32426	Translated from 5'
9	101	100.0	2227	AAW05697	Attenuated hepatit
10	101	100.0	2227	AAW34074	Hepatitis A virus

11	101	100.0	2227	21	AAW18607	Amino acid sequenc
12	101	100.0	2227	21	AAW18608	Amino acid sequenc
13	101	100.0	2227	21	AAW18609	Amino acid sequenc
14	101	100.0	2227	23	AAW18609	Hepatitis A virus
15	96	95.0	2227	7	AAW60066	Sequence of viral
16	84	83.2	839	12	AAW15629	Capsid region of c
17	65	64.4	25	18	AAW42930	Immunogenic Hepati
18	65	64.4	26	22	AAW69447	Synthetic HAV P2A
19	52.5	52.0	1025	16	AAW86001	Human dihydroxyrim
20	52.5	52.0	1025	17	AAW86001	Human dihydroxyrim
21	52.5	52.0	1025	20	AAW93361	Human DPD protein.
22	47	46.5	407	21	AAW06085	Arabidopsis thalia
23	46	45.5	117	22	AAW00818	Human polypeptide
24	45.5	45.0	114	22	AAW03241	Human polypeptide
25	45.5	45.0	1025	16	AAW86000	Bovine dihydroxyri
26	45	44.6	23	12	AAW15517	BMP-8 peptide. Bo
27	45	44.6	119	21	AAW05136	Arabidopsis thalia
28	45	44.6	119	21	AAW49796	Arabidopsis thalia
29	45	44.6	132	21	AAW05135	Arabidopsis thalia
30	45	44.6	132	21	AAW49795	Arabidopsis thalia
31	45	44.6	150	21	AAW05134	Arabidopsis thalia
32	45	44.6	150	21	AAW36068	Zea mays protein f
33	45	44.6	150	21	AAW49794	Arabidopsis thalia
34	45	44.6	178	21	AAW05133	Arabidopsis thalia
35	45	44.6	178	21	AAW36067	Zea mays protein f
36	45	44.6	178	21	AAW49793	Arabidopsis thalia
37	45	44.6	517	23	AAW74627	Oestrogen-regulate
38	45	44.6	519	22	ABW71162	Drosophila melanog
39	45	44.6	519	23	AAW74628	Oestrogen-regulate
40	44	43.6	65	22	AAW12217	Human polypeptide
41	44	43.6	187	21	AAW32794	Eucalyptus grandis
42	44	43.6	445	22	AAW82882	S. epidermidis ope
43	44	43.6	454	23	ABW39593	Staphylococcus epi
44	44	43.6	673	23	AAW99901	Human 47476 guanin
45	43.5	43.1	97	22	ABW50902	Human secreted pro

ALIGNMENTS

RESULT 1
AAW42925
ID AAW42925 standard; peptide; 20 AA.
XX
AC AAW42925;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1318.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX Claim 18; Page 112; 140pp; English.
PS
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 823-842, and has a reactivity of 31.3% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
XX Sequence 20 AA;
Query Match 100.0%; Score 101; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 VLPPPRKMKGLFSQAKISLF 20
1 VLPPPRKMKGLFSQAKISLF 20
1 VLPPPRKMKGLFSQAKISLF 20
RESULT 2
AAB69442
ID AAB69442 standard; Peptide; 21 AA.
XX
AC AAB69442;
XX
DT 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 42.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
PR 15-JUL-1999; 99US-0144412.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 95; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IGM antibody reactivity.
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 101; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
1 VLPPPRKMKGLFSQAKISLF 20
1 VLPPPRKMKGLFSQAKISLF 20
Db 1 VLPPPRKMKGLFSQAKISLF 20
1 VLPPPRKMKGLFSQAKISLF 20
RESULT 3
AAP50230
ID AAP50230 standard; Protein; 366 AA.
XX
AC AAP50230;
XX
DT 28-NOV-1991 (first entry)
XX
DE Sequence of hepatitis A virus (HAV) surface protein (VP-I).
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
OS Hepatitis A virus.
XX
PN EPI38704-A.
XX
PD 24-APR-1985.
XX
PF 09-OCT-1984; 84EP-0402025.
XX
PR 02-MAR-1984; 84US-0585942.
XX 14-OCT-1983; 83US-0541836.
XX
PA (MERI) MERCK & CO INC.
XX
PI Hughes JV, Scolnick EM, Tomassini JE;
XX
DR WPI; 1985-100818/17.
XX
DR N-PSDB; AAN50274.
XX
PT New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus
XX
PS Claim 21; Page 46-48; 49pp; English.
XX
CC VP1 is isolated by solubilisation of the intact virus in an aq.
CC anionic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ Sequence 366 AA;
Query Match 100.0%; Score 101; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
332 VLPPPRKMKGLFSQAKISLF 351
Db 332 VLPPPRKMKGLFSQAKISLF 351
RESULT 4
AAP50287
ID AAP50287 standard; Protein; 854 AA.
XX
AC AAP50287;
XX
DT 30-NOV-1991 (first entry)

```

XX WPI: 1985-224964/37.
DR N-PSDB; AAN50139.
XX
XX New nucleotide sequences coding for hepatitis A virus antigens -
PT useful for eliciting normal immune response and in vaccines for
PT protecting against the virus
XX
XX Example; Page 11-17; 32pp; English.
XX
XX Within the sequence in AAN50139 is encoded the information necessary
CC to make the antigenic proteins of HAV. The sequences encoding for
CC the structural proteins begin at base 403. The key sub-unit
CC sequences within VP-1, designated Sequences I,II,III,IV, and V,
CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
CC nucleotide sequences which are valuable as encoding antigenic
CC proteins are the sequences from base 1749 to base 2722; from base
CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
CC base 1749 to base 2722 is esp. valuable as a vector for producing
CC antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
CC translation of a stop codon.
XX
XX Sequence 993 AA;
XX
XX Query Match 100.08; Score 101; DB 6; Length 993;
XX Best Local Similarity 100.08; Pred. No. 1.5e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Caps
XX
XX QY 1 VLP PPRKMKGLFSQAKISLF 20
XX |||||||||||||||||||
XX DB 959 VLP PPRKMKGLFSQAKISLF 978
XX
XX RESULT 6
XX AAP50231
XX ID AAP50231 standard; Protein; 993 AA.
XX
XX AAP50231;
XX
XX 28-NOV-1991 (first entry)
XX
XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
DE
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX diagnostic assay.
XX
XX Hepatitis A virus.
XX
XX Key Location/Qualifiers
XX FH 628..993
XX Protein /note= "claimed; X denotes translated stop codons
XX FT and unspecified triplets"
XX
XX EPI38704-A.
XX
XX 24-APR-1985.
XX
XX 09-OCT-1984; 84EP-0402025.
XX
XX 02-MAR-1984; 84US-0585942.
XX 14-OCT-1983; 83US-0541836.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
XX
XX WPI: 1985-100818/17.
XX N-PSDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus
XX

```


CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (I) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (I); (2) a cell transfected with (I) or the RNA
CC transcript of (I); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (I) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (I) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
|||||
DB 823 VLPPrKMKGLFSQAKISLF 842

RESULT 11
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
XX
AC AAB18607;
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
XX
PT 17-SEP-1993; 93WO-US08610.
XX
PT 10-MAR-1995; 95US-0397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI; 2000-586464/55.
XX
DR N-PSDB; AAA75476.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
PS Disclosure; Fig 6A-K; 72pp; English.
XX
CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
|||||
DB 823 VLPPrKMKGLFSQAKISLF 842

RESULT 12
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
XX
AC AAB18608;
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW P-35 virus.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
XX
PT 17-SEP-1993; 93WO-US08610.
XX
PT 10-MAR-1995; 95US-0397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI; 2000-586464/55.
XX
DR N-PSDB; AAA75477.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
PS Disclosure; Columns 67-78; 72pp; English.
XX
CC The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
|||||
DB 823 VLPPrKMKGLFSQAKISLF 842

RESULT 13
AAB18609
ID AAB18609 standard; Protein; 2227 AA.
XX
AC AAB18609;
DT 15-JAN-2001 (first entry)

Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
|||||
DB 823 VLPPrKMKGLFSQAKISLF 842

RESULT 13
AAB18609
ID AAB18609 standard; Protein; 2227 AA.
XX
AC AAB18609;
DT 15-JAN-2001 (first entry)


```

XX 29-OCT-1986.
XX
XX 03-APR-1986; 86EP-0302465.
XX
XX 03-APR-1985; 85US-0719329.
XX
XX (CHIR-) CHIRON CORP.
XX
XX Dina D, Potter SJ, Vannest GA, Caput D;
XX
XX WPI; 1986-286213/44.
XX
XX N-PSDB; AAN60080.
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use
XX in prodn. of vaccines and diagnostic probes
XX
XX Claim 5; Fig 1; 18pp; English.
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
XX hepatitis A virus; transformed hosts may be used for expression of
XX polypeptides and fragments useful in vaccines without risk of
XX infection by the virus or in prodn. of particles which are capable
XX of inducing immunocompetent B cells for passive immunotherapy. Pref.
XX epitope is derived from AAs 445-657 or 792-848 of the HAV
XX polypeptide sequence (AAP60066).
XX
XX Sequence 2227 AA;
XX
Query Match 95.0%; Score 96; DB 7; Length 2227;
Best Local Similarity 95.0%; Pred. NO. 2.le-06;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842
Search completed: February 3, 2003, 13:11:30
Job time : 29.8333 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 : Search time 9.5 Seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6	Patent No. 5516630-2
2	101	100.0	2227	3	US-08-475-886-2
3	101	100.0	2227	3	US-08-475-886-4
4	101	100.0	2227	3	US-08-475-886-6
5	101	100.0	2227	4	US-08-397-232-2
6	101	100.0	2227	4	US-08-397-232-4
7	101	100.0	2227	4	US-09-171-387-2
8	101	100.0	2227	4	US-09-653-499-2
9	101	100.0	2227	4	US-09-653-499-4
10	101	100.0	2227	4	US-09-653-499-6
11	84	83.2	839	1	US-08-087-016-2
12	52.5	52.0	1025	3	US-08-304-309-2
13	52.5	52.0	1025	3	US-08-991-942-2
14	52.5	52.0	1025	4	US-09-138-103-2
15	52.5	52.0	1025	5	PCT-US95-04567-4
16	45.5	45.0	1025	5	PCT-US95-04567-2
17	45	44.6	23	1	US-07-800-364B-4
18	45	44.6	23	5	PCT-US91-03388-4
19	44	43.6	454	4	US-09-134-001C-4438
20	41	40.6	108	1	US-08-466-033-106
21	41	40.6	108	2	US-08-444-733-106
22	41	40.6	108	2	US-08-464-134-106
23	41	40.6	108	2	US-08-461-361-106
24	41	40.6	108	2	US-08-485-910-106
25	41	40.6	108	5	PCT-US95-06266-89
26	41	40.6	616	1	US-08-638-911A-35
27	41	40.6	853	1	US-08-638-911A-25

28 41 40.6 853 1 US-08-638-911A-27
29 41 40.6 853 1 US-08-638-911A-29
30 41 40.6 853 1 US-08-638-911A-31
31 41 40.6 2873 1 US-08-466-033-15
32 41 40.6 2873 1 US-08-638-911A-2
33 41 40.6 2873 2 US-08-444-733-15
34 41 40.6 2873 2 US-08-464-134-15
35 41 40.6 2873 2 US-08-461-361-15
36 41 40.6 2873 2 US-08-485-910-15
37 41 40.6 2873 5 PCT-US95-06266-15
38 41 40.6 2905 4 US-08-469-260A-401
39 41 40.6 2910 1 US-08-466-033-183
40 41 40.6 2910 2 US-08-444-733-183
41 41 40.6 2910 2 US-08-464-134-183
42 41 40.6 2910 2 US-08-461-361-183
43 41 40.6 2910 2 US-08-485-910-183
44 41 40.6 2910 5 PCT-US95-06266-157
45 40 39.6 34 2 US-08-751-767A-38

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 1060 VLPPPRKMKGLFSQAKISLF 1079

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475.886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

SULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match      100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | |
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 7

US-09-171-387-2

; Sequence 2, Application US/09171387

; Patent No. 6280734

; GENERAL INFORMATION:

; APPLICANT: RAYCHAUDHURI, GOPA;

; EMERSON, SUZANNE, U.;

; PURCELL, ROBERT, H.

; TITLE OF INVENTION: SIMIAN-HUMAN HAV

; HAVING A CHIMERIC 2C PROTEIN

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: MICROSOFT WORD 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/171.387

; FILING DATE: 24-Mar-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US97/06506

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: US60/015,642

; FILING DATE: 19-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: William S. Feiler

; REGISTRATION NUMBER: 26,728

; REFERENCE/DOCKET NUMBER: 2026-4229US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 2

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2227 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-171-387-2

Query Match 100.0%; Score 101; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | |
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 8

US-09-653-499-2

; Sequence 2, Application US/09653499

; Patent No. 6423318

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/09/653,499

; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | |
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 9

US-09-653-499-4

; Sequence 4, Application US/09653499

; Patent No. 6423318

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/09/653,499

; CURRENT FILING DATE: 2000-08-31

; PRIOR APPLICATION NUMBER: 08/475,886

; PRIOR FILING DATE: 1999-08-09

; PRIOR APPLICATION NUMBER: 07/947,338

; PRIOR FILING DATE: 1992-09-18

; PRIOR APPLICATION NUMBER: 08/397,232

; PRIOR FILING DATE: 1995-03-10

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 2227

; TYPE: PRT

; ORGANISM: Attenuated HAV (Pass 35), strain HM-175

US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
| | | | | | | | | | | | | | | | | |
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 10

US-09-653-499-6

; Sequence 6, Application US/09653499

; Patent No. 6423318

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/09/653,499

; CURRENT FILING DATE: 2000-08-31

;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||:|||||:|||||

823 VLPPPRKMKGLFSQAKISLF 842

RESULT 11

US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087.016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 4.3e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKIS 18
:|||||:|||||:|||||

DB 822 ILPPPRKMKGLFSQSKIS 839

RESULT 12

US-08-304-309-2
; Sequence 2, Application US/08304309
; Patent No. 5856454
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-304-309-2

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.1;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGLFSQAKISLF 20
|||||:|||||:|||||

DB 176 LPPPEKMESEYS-AKIALF 193

RESULT 13

US-08-991-942-2
; Sequence 2, Application US/08991942
; Patent No. 6015673
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493


```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/991,942
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-942-2

Query Match 52.0%; Score 52.5; DB 3; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.1;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Oy 2 LPPPRKMKGLFSQAKISLF 20
    |||| || : |||:|
Db 176 LPPPEKMEAYS-AKIALF 193

RESULT 14
US-09-138-103-2
; Sequence 2, Application US/09138103A
; Patent No. 6232448
; GENERAL INFORMATION:
; APPLICANT: Yoshikubo, Takashi
; APPLICANT: Hasegawa, Masami
; TITLE OF INVENTION: Immunological Materials and Methods for Detecting
; TITLE OF INVENTION: Dihydropyrimidine Dehydrogenase
; FILE REFERENCE: 09/138,103 Yoshikubo, et al.
; CURRENT APPLICATION NUMBER: US/09/138,103A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 97114630.3
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-138-103-2

Query Match 52.0%; Score 52.5; DB 4; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.1;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Oy 2 LPPPRKMKGLFSQAKISLF 20
    |||| || : |||:|
Db 176 LPPPEKMEAYS-AKIALF 193

RESULT 15
PCT-US95-04567-4
; Sequence 4, Application PC/TUS9504567
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
```

```
;
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04567
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,357
; FILING DATE: 13-APR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UOAB025P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04567-4

Query Match 52.0%; Score 52.5; DB 5; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.1;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Oy 2 LPPPRKMKGLFSQAKISLF 20
    |||| || : |||:|
Db 176 LPPPEKMEAYS-AKIALF 193

Search completed: February 3, 2003, 13:16:38
Job time : 11.5 secs
```

THIS PAGE BLANK (USPTO)


```
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842
|||||
```

```
RESULT 3
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842
|||||
```

```
RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
```

```
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842
|||||
```

```
RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
```

```
Query Match 100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842
|||||
```

```
RESULT 6
US-09-854-886-2
; Sequence 2, Application US/09854886
; Patent No. US20020072080A1
; GENERAL INFORMATION:
; APPLICANT: Yoshikubo, Takashi
; APPLICANT: Hasegawa, Masami
; TITLE OF INVENTION: Immunological Material and Methods for Detecting
; TITLE OF INVENTION: Dihydropyrimidine Dehydrogenase
; FILE REFERENCE: 100554-32887
; CURRENT APPLICATION NUMBER: US/09/854,886
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/138,103
; PRIOR FILING DATE: 1998-08-21
```

; PRIOR APPLICATION NUMBER: 97114630.3
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-886-2

Query Match 52.0%; Score 52.5; DB 10; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.8;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Oy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPEKMESEYS-AKIALF 193
|||||||:|||||

RESULT 7

US-10-017-216-2
; Sequence 2, Application US/10017216
; Patent No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPPELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2

Query Match 41.6%; Score 42; DB 9; Length 2053;
Best Local Similarity 61.5%; Pred. No. 1.9e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 4 PPRKMKGLFSQAK 16
Db 1259 PAKKKKGLFSRKK 1271
|:|||||:

RESULT 8

US-08-424-550B-401
; Sequence 401, Application US/08424550B
; Patent No. US20020119447A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA

ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 401:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-401

Query Match 40.6%; Score 41; DB 8; Length 2905;
Best Local Similarity 46.7%; Pred. No. 4.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 3 PPRKMKGLFSQAKI 17
Db 1140 PPPVPKGVKEAPL 1154
||| |||:

RESULT 9

US-09-851-026-11
; Sequence 11, Application US/09851026
; Patent No. US20020160957A1
; GENERAL INFORMATION:
; APPLICANT: Stern, David M.
; Clauss, Matthias
; Kao, Janet
; Kayton, Mark
; Libutti, Steven K
; TITLE OF INVENTION: Endothelial Monocyte Activating
; Polypeptide II: A Mediator Which Activates Host Respons
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham, LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,026
; FILING DATE: 07-May-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,821
; FILING DATE: 08-OCT-96
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525

```
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 166 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-851-026-11

Query Match          39.6%; Score 40; DB 9; Length 166;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 69 PAKMRGVLSQAMV 81

RESULT 10
US-09-930-169-3
Sequence 3, Application US/0993010169
Publication No. US20030004309A1
GENERAL INFORMATION:
APPLICANT: KIM, SUNGHOON
TITLE OF INVENTION: IMMUNOLOGICAL ENHANCEMENT AGENT COMPRISING N-TERMINAL PEPTIDE OF
FILE REFERENCE: 058333/0106
CURRENT APPLICATION NUMBER: US/09/930,169
CURRENT FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: KR 2001-31310
PRIOR FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 166
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Mammalian
OTHER INFORMATION: Protein sequence
US-09-930-169-3

Query Match          39.6%; Score 40; DB 9; Length 166;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 125 PAKMRGVLSQAMV 137

RESULT 11
US-09-851-026-35
Sequence 35, Application US/09851026
Patent No. US20020160957A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
Claus, Matthias
Kao, Janet
Kayton, Mark
Libutti, Steven K
TITLE OF INVENTION: Endothelial Monocyte Activating
Polyptide II: A Mediator Which Activates Host Response
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

Query Match          39.6%; Score 40; DB 9; Length 166;
Best Local Similarity 53.8%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 213 PAKMRGVLSQAMV 225

RESULT 12
US-09-851-026-36
Sequence 36, Application US/09851026
Patent No. US20020160957A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
Claus, Matthias
Kao, Janet
Kayton, Mark
Libutti, Steven K
TITLE OF INVENTION: Endothelial Monocyte Activating
Polyptide II: A Mediator Which Activates Host Response
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

Query Match          39.6%; Score 40; DB 9; Length 310;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 213 PAKMRGVLSQAMV 225
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.30, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/851,026
; FILING DATE: 07-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,821
; FILING DATE: 08-OCT-96
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 310 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-09-851-026-35

Query Match          39.6%; Score 40; DB 9; Length 310;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 213 PAKMRGVLSQAMV 225

RESULT 12
US-09-851-026-36
Sequence 36, Application US/09851026
Patent No. US20020160957A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
Claus, Matthias
Kao, Janet
Kayton, Mark
Libutti, Steven K
TITLE OF INVENTION: Endothelial Monocyte Activating
Polyptide II: A Mediator Which Activates Host Response
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

Query Match          39.6%; Score 40; DB 9; Length 310;
Best Local Similarity 53.8%; Pred. No. 44;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 PRKMKGLFSQAKI 17
Db 213 PAKMRGVLSQAMV 225
```

; REFERENCE/DOCKET NUMBER: 41735-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-851-026-36

Query Match 39.6%; Score 40; DB 9; Length 312;
Best Local Similarity 53.8%; Pred. No. 45;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 5 PRKMKGLFSQAKI 17
Db 215 PAKMRGVLQAMV 227

RESULT 13
US-09-947-971-2
; Sequence 2, Application US/09947971
; Patent No. US20020143148A1
; GENERAL INFORMATION:
; APPLICANT: Lok, SI
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Human Vomeronasal Receptor-4
; FILE REFERENCE: 00-83
; CURRENT APPLICATION NUMBER: US/09/947,971
; CURRENT FILING DATE: 2001-09-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-971-2

Query Match 39.6%; Score 40; DB 10; Length 318;
Best Local Similarity 64.3%; Pred. No. 46;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 6 RKMKGFLFSQAKISL 19
Db 19 QKTRNLFQATIGL 32

RESULT 14
US-09-320-337-76
; Sequence 76, Application US/09320337
; Patent No. US20010016352A1
; GENERAL INFORMATION:
; APPLICANT: Bohinski, Robert J.,
; APPLICANT: Whitsett, Jeffrey A.
; TITLE OF INVENTION: Nucleic Acid Sequences Controlling
; TITLE OF INVENTION: Lung Cell - Specific Gene Expression
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM P160
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: MS WORD 97

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/320,337
; FILING DATE: 26-MAY-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/442,809
; FILING DATE: 17-MAY-1995
; APPLICATION NUMBER: 08/245,356
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Olstein, Elliot M.
; REGISTRATION NUMBER: 24,025
; REFERENCE/DOCKET NUMBER: 271010-447
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 973-994-1700
; TELEFAX: 973-994-1744
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: human thyroid transcription factor-1
US-09-320-337-76

Query Match 39.6%; Score 40; DB 10; Length 371;
Best Local Similarity 57.1%; Pred. No. 55;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy 4 PPRKMKGLFSQAKI 17
Db 160 PRRKRRVLSQAV 173

RESULT 15
US-09-815-242-4933
; Sequence 4933, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4933
; LENGTH: 666
; TYPE: PRT

; ORGANISM: Enterococcus faecalis
US-09-815-242-4933

Query Match 39.6%; Score 40; DB 10; Length 666;
Best Local Similarity 61.5%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKI 17
||||:| || |
Db 629 PRKMRGQISQGM I 641

Search completed: February 3, 2003, 13:32:25
Job time : 7 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 ; Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	852	1 GNNVHA	genome polyprotein
2	101	100.0	1358	2 A03905	genome polyprotein
3	101	100.0	2227	1 GNNVHM	genome polyprotein
4	101	100.0	2227	1 GNNVHR	genome polyprotein
5	101	100.0	2227	1 GNNYMK	genome polyprotein
6	96	95.0	2230	1 GNNYSA	genome polyprotein
7	95	94.1	2227	1 GNNYHB	genome polyprotein
8	84	83.2	839	1 GNNYS2	genome polyprotein
9	75	74.3	341	2 S04137	genome polyprotein
10	52.5	52.0	1025	2 A54718	dihydropyrimidine
11	47	46.5	414	2 D96838	unknown protein T2
12	45	44.6	79	2 AG1062	bacteriophage gene
13	45	44.6	321	2 E86423	probable 60S ribos
14	44	43.6	340	1 WMBES7	ribonucleoside-dip
15	44	43.6	343	2 S74937	hypothetical prote
16	44	43.6	369	2 S77299	hypothetical prote
17	44	43.6	775	2 T48957	hypothetical prote
18	43	42.6	189	2 A12534	transposase all1756
19	43	42.6	309	2 D75008	methanol dehydroge
20	43	42.6	379	2 T49919	hypothetical prote
21	43	42.6	389	2 S41748	heat shock protein
22	42.5	42.1	105	2 B72598	hypothetical prote
23	42	41.6	109	2 S69307	probable membrane
24	42	41.6	165	2 S02053	hypothetical prote
25	42	41.6	178	2 T02335	ribosomal protein
26	42	41.6	834	2 F82673	hypothetical prote
27	41.5	41.1	304	2 H82984	hypothetical prote
28	41	40.6	83	2 T17809	hypothetical prote
29	41	40.6	159	2 C72210	conserved hypothet

30	41	40.6	213	2 H86638	hypothetical prote
31	41	40.6	229	2 T19153	hypothetical prote
32	41	40.6	237	2 C84065	transcription regu
33	41	40.6	370	2 G97001	endoglucanase faml
34	41	40.6	561	2 G96752	unknown protein F2
35	41	40.6	622	2 T37257	hypothetical prote
36	41	40.6	897	2 T06540	polyribonucleotide
37	41	40.6	1040	2 A34695	axonal glycoprotein
38	41	40.6	1234	2 T30254	jumonji protein -
39	40.5	40.1	924	2 T25007	hypothetical prote
40	40.5	40.1	932	2 T25008	hypothetical prote
41	40	39.6	177	2 C64133	mercuric resistanc
42	40	39.6	238	2 C83240	probable transcrip
43	40	39.6	259	2 F64532	conserved hypothet
44	40	39.6	272	2 T49070	probable heat choc
45	40	39.6	292	2 S23239	hypothetical prote

ALIGNMENTS

RESULT 1

GNNVHA

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03904
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra,
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793
A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; core protein; polyprotein

F;1-245/Product: coat protein 1A #status predicted <CIA>

F;246-491/Product: coat protein 1B #status predicted <CLB>

F;492-836/Product: coat protein 1C #status predicted <CLC>

F;837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 101; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

|||||

Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 2

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Tichehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Felns

proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; core protein; polyprotein

F;1-245/Product: coat protein 1A #status predicted <CIA>

F;246-491/Product: coat protein 1B #status predicted <CLB>

F;492-836/Product: coat protein 1C #status predicted <CLC>

F;837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 101; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||

Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 3

GNMYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 3D

B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Naajarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127

C:Accession: A03903

A:Molecule type: genomic RNA

A:Residues: 1-2227 <NAJ>

A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||

Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 4

GNMYHM

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro

NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03903

R:Naajarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127

C:Accession: A03903

A:Molecule type: genomic RNA

A:Residues: 1-2227 <NAJ>

A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||

Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5

GNMYMK

genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core

NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999

C:Accession: A94149; A25914; A94508

R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell

Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987

A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison

A:Reference number: A94149; MUID:87175701; PMID:3031686

A:Accession: A94149

A:Status: nucleic acid sequence not shown

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595

A:Note: submitted to Genbank, August 1987

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-245/Product: coat protein 1A #status predicted <P1A>

F:246-491/Product: coat protein 1B #status predicted <P1B>

F:492-836/Product: coat protein 1C #status predicted <P1C>

F:837-980/Product: core protein 2A #status predicted <P2A>

F:981-1076/Product: core protein 2B #status predicted <P2B>

F:1077-1422/Product: core protein 2C #status predicted <P2C>

F:1423-1484/Product: protein 3A #status predicted <P3A>

F:1485-1507/Product: protein 3B #status predicted <P3B>

F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>

F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||

Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 6

GNMYSA

genome polyprotein - simian hepatitis A virus (strain AGM-27)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core

C:Species: simian hepatitis A virus

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000

C:Accession: A30470; S04885; S03965

R:Tzarev, S.A.

submitted to JIPID, April 1991

A:Reference number: A30470

A:Accession: A30470

A:Molecule type: genomic RNA

A:Residues: 1-2230 <TSA>

A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598

R:Tzarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.

J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
A:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 95.0%; Score 96; DB 1; Length 2230;
Best Local Similarity 90.0%; Pred. No. 2.5e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
:|||||:|||||:|||||
Db 827 ILPPPRKMKGLFSQAKISLF 846

RESULT 7
GNNYHB
genome polyprotein - human hepatitis A virus (strain MBB)
N:Contents: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
;Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt,
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyd
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP4>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 94.1%; Score 95; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 3.6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
:|||||:|||||:|||||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 8
GNNYS2
genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
N:Contents: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat prot
C:Species: simian hepatitis A virus
A:Note: host Macaca fascicularis (cynomolgus macaque)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: J01180
R:Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
J. Gen. Virol. 72, 1685-1689, 1991
A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus
A:Reference number: J01180; MUID:91311421; PMID:1649902
A:Accession: J01180
A:Molecule type: genomic RNA
A:Residues: 1-839 <NAI>
A:Cross-references: GB:M59286; NID:g329599; PIDN:AAA45473.1; PID:g555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 8.5e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKIS 18
:|||||:|||||:|||||
Db 822 ILPPPRKMKGLFSQAKIS 839

RESULT 9
S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir
A:Reference number: S04137; MUID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 74.3%; Score 75; DB 2; Length 341;
Best Local Similarity 93.3%; Pred. No. 0.0001;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQA 15
:|||||:|||||:|||||
Db 327 VLPPPRKMKGLFSQS 341

RESULT 10

A54718
dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 03-Jun-2002
C:Accession: A54718
R:Yokota, H.; Fernandez-Salguero, P.; Furuya, H.; Lin, K.; McBride, O.W.; Podschun, B.; J. Biol. Chem. 269, 23192-23196, 1994
A:Title: cDNA cloning and chromosome mapping of human dihydropyrimidine dehydrogenase, a
A:Reference number: A54718; MUID:94365020; PMID:8083224
A:Accession: A54718
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1025 <YOK>
A:Cross-references: GB:U09178; NID:g558304; PID:g558305
A:Note: authors translated the codon GGC for residue 748 as Asp, and TTA for residue 101
C:Genetics:
A:Gene: GDB:DPYD
A:Cross-references: GDB:364102; OMIM:274270
A:Map position: lp22-lp22
A:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
Keywords: 4Fe-4S; flavoprotein; homodimer; iron-sulfur protein; NADP; oxidoreductase
946-1005/Domain: ferredoxin 2[4Fe-4S] homology <FER>
Query Match 52.0%; Score 52.5; DB 2; Length 1025;
Best Local Similarity 63.2%; Pred. No. 2;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 2 LPPPRKMKGLFSQAKISLF 20
||||| : : : : :
Db 176 LPPPRKMKSEAYS-AKIALF 193
RESULT 11
D96838
unknown protein T21F11.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96838
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewart, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzlia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96838
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <SOT>
A:Cross-references: GB:AE005173; NID:96730725; PIDN:AAF27115.1; GSPDB:GN00141
C:Genetics:
A:Gene: T21F11.5
A:Map position: 1
Query Match 46.5%; Score 47; DB 2; Length 414;
Best Local Similarity 56.2%; Pred. No. 6.1;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 3 PPRPRKMKGLFSQAKIS 18
||||| : : : : :
Db 106 PPRPRKMKGLFSQAKIS 121
RESULT 12
AG1062
bacteriophage gene regulatory protein STV4826 [imported] - Salmonella enterica subsp. en
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C:Accession: AG1062
C:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
A:Reference number: AB0502; PMID:11677608
A:Accession: AG1062
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-79 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06948.1; PID:g16505595; GSPDB:GN00176
C:Genetics:
A:Gene: STV4826
Query Match 44.6%; Score 45; DB 2; Length 79;
Best Local Similarity 63.6%; Pred. No. 2.1;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 3 PPRPRKMKGLFS 13
||||| : : : : :
Db 64 PPRPRKMQCYS 74
RESULT 13
E86423
probable 60S ribosomal protein L18A - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: E86423
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewart, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzlia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86423
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <SOT>
A:Cross-references: GB:AE005172; NID:g10092460; PIDN:AAG12862.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
Query Match 44.6%; Score 45; DB 2; Length 321;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
Qy 1 VLPPrKMKGLFSQAKISLF 20
| | | | | : : : : :
Db 301 VRPPrKLTTFKANKPNLF 320
RESULT 14
WMBES7
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human herpesvirus 1
N:Alternate names: ribonucleotide reductase small chain
C:Species: human herpesvirus 1
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: D30088
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; P
J. Gen. Virol. 69, 1531-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes
A:Reference number: A30083; MUID:88274327; PMID:2839594
A:Accession: D30088
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A:Residues: 1-340 <MCG>
A:Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32303.1; PID:g59540; GB:D00317
C:Genetics:
A:Gene: UL40
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain
C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; oxidoreductase
F:94,124,127,187,221,224/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #stable
F:131/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 43.6%; Score 44; DB 1; Length 340;
Best Local Similarity 61.5%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPRKMKGLFSQA 15
||| : ||| ||
Db 242 PPDPRVYGLFRQA 254

RESULT 15
S74937
hypothetical protein slr0688 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S74937
R:Kansko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O., K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74937
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <KAN>
A:Cross-references: EMBL:D90902; GB:AB001339; NID:g1652027; PIDN:BAAL6977.1; PID:dl01771
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 43.6%; Score 44; DB 2; Length 343;
Best Local Similarity 47.1%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPRKMKGLFSQAKISL 19
| ||: |||| |:::
Db 240 PDPRECKGLFRGAFTI 256

Search completed: February 3, 2003, 13:15:34
Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
1	101	100.0	852	1	POLG_HPAPVC	P06442 hepatitis a	
2	101	100.0	852	1	POLG_HPAPVC	P08617 hepatitis a	
3	101	100.0	2227	1	POLG_HPAPV2	P06441 hepatitis a	
4	96	95.0	2226	1	POLG_HPAPV2	P26580 hepatitis a	
5	96	95.0	2226	1	POLG_HPAPV8	P26581 hepatitis a	
6	96	95.0	2226	1	POLG_HPAPV8	P26582 hepatitis a	
7	96	95.0	2230	1	POLG_HPAPV5	P14553 simian hepa	
8	95	94.1	2227	1	POLG_HPAPV7	P13901 hepatitis a	
9	84	83.2	839	1	POLG_HPAPV7	P13788 simian hepa	
10	75	74.3	341	1	POLG_HPAPV1	P13672 hepatitis a	
11	52.5	52.0	1025	1	DPYD_HUMAN	Q28862 homo sapien	
12	45.5	45.0	1025	1	DPYD_BOVIN	Q28007 bos taurus	
13	45	44.6	178	1	RL1X_ORYSA	Q943f3 oryza sativ	
14	45	44.6	178	1	RL1X_ORYSA	Q91ud4 arabidopsis	
15	44	43.6	340	1	R1R2_HSV11	P10224 herpes simp	
16	44	43.6	340	1	R1R2_HSV1K	P06474 herpes simp	
17	43	42.6	389	1	DNAJ_METWA	P35515 methanosarc	
18	43	42.6	587	1	MTSL_STRAL	Q53609 streptomyce	
19	42	41.6	178	1	RL1X_ORYSA	P51418 arabidopsis	
20	41	40.6	289	1	NR2F_MOUSE	P43688 mus musculu	
21	41	40.6	359	1	MCAL_CRIGR	O54873 cricetus	
22	41	40.6	1040	1	AXOL_RAT	P22063 rattus norv	
23	41	40.6	1234	1	JMJ_MOUSE	Q62315 mus musculu	
24	40	39.6	135	1	ZNTR_HAETIN	P45277 haemophilus	
25	40	39.6	292	1	YOMI_CAEL	P30647 caenorhabdi	
26	40	39.6	310	1	MCAL_MOUSE	P31230 mus musculu	
27	40	39.6	312	1	MCAL_HUMAN	Q12904 homo sapien	
28	40	39.6	371	1	TTFL_CANFA	P43698 canis famil	
29	40	39.6	371	1	TTFL_HUMAN	P43699 homo sapien	
30	40	39.6	372	1	TTFL_MOUSE	P50220 mus musculu	
31	40	39.6	372	1	TTFL_RAT	P23441 rattus norv	
32	40	39.6	448	1	DHE4_HELPY	P55990 helicobacte	
33	40	39.6	523	1	YMR9_YEAST	Q05040 saccharomyc	

RESULT 1

POLG_HP4VC	STANDARD:	PRT:	852 AA.
OC	POLG_HP4VC		
AD	P063741; Q83742;		
DT	01-JAN-1988 (Rel. 06, Created)		
DT	01-JAN-1988 (Rel. 06, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DT	Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core protein		
DE	p2A] (Fragment).		
OC	Hepatitis A virus (strain CR326).		
OC	Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
OC	Hepatovirus.		
OC	NCBI_TaxID=12097;		
RN	[1]		

ALIGNMENTS

Query Match	100.0%	Score 101:	DB 1:	Length 852:
Best Local Similarity	100.0%	Pred. NO. 6.9e-09;		
Matches 20:	Conservative	0:	Mismatches 0:	Indels 0:
				Gaps 0:

RESULT 2

РЕЗУЛТАТИ
ПОЛГ ПРАВН

ID POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P06617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RA "Complete nucleotide sequence of wild-type hepatitis A virus:
RA comparison with different strains of hepatitis A virus and other
RA picornaviruses.";
RA J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,
RA Purcell R.H.;
RA "Complete nucleotide sequence of an attenuated hepatitis A virus:
RA comparison with wild-type virus.";
RA Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstein S.M.;
RA "Sequence analysis of hepatitis A virus cDNA coding for capsid
RA proteins and RNA polymerase.";
RA Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M14114; AAA45475.1; -.
DR EMBL; M14707; AAA45465.1; -.
DR EMBL; M14707; AAA45466.1; ALT_INIT.
DR EMBL; M16632; AAA45471.1; -.
DR PIR; A25981; GNNYHM.
DR PIR; A25914; GNNYMK.
DR PIR; A03905; A03905.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7A5B740A6 CRC64;
Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 3
POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dina D.;
RA "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; K02990; AAA45472.1; -.
DR PIR; A03903; GNNYHR.


```
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
   |||||
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 4
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59810; AAA45468.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
```

```
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D684E2BF CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
   |||||
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5
POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59809; AAA45469.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
```

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4C8A80B9BF75 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 6
POLG_HPAV8
ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Croneans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59808; AAA45467.1; -;
DR MEROPS; C03.005; -;
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A6396C8D6B CRC64;
Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 7
POLG_HPAV8
ID POLG_HPAV8 STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00924; BAA00766.1; -;
DR EMBL; X15461; CAA33490.1; -;
DR Pfam; A30470; GNNYSA.
DR Pfam; S04885; S04885.
DR MEROPS; C03.005; -;
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2230;
Best Local Similarity 90.0%; Pred. No. 1.3e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLPPrKMKGLFSQAKISLF 20
:|||||:|||||
b 827 ILPPPRKVKGLFSQAKISLF 846

RESULT 8
POLG_HPAVM STANDARD; PRT; 2227 AA.
ID AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MB8).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wissel T., Klehn R., Wimmer E.,
Reinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MB8).";
RL Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
[RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M20273; AAA45474.1; -
DR PIR; J50303; GNNYHB.
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00660; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
FT CHAIN 837 980 CORE PROTEIN P2A.
FT CHAIN 981 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 94.1%; Score 95; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 1.9e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLPPrKMKGLFSQAKISLF 20
:|||||:|||||
Db 823 VLPPrKMKGLFSQAKISLF 842

RESULT 9
POLG_HPAVT STANDARD; PRT; 839 AA.
ID AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
P2A] (Fragment).
DE Simian hepatitis A virus (strain CV-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59286; AAA45473.1; -
DR PIR; JQ1180; GNNYS2.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACCB0D1E192DBC CRC64;
Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 4.4e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLPPrKMKGLFSQAKIS 18
:|||||:|||||
Db 822 ILPPPrKMKGLFSQAKIS 839

```
RESULT 10
POLG_HPAV1
ID POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 to VP3; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain LCDG-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Anonov A.P., Lau P., Chaudhary R.;
RA "Nucleotide sequence of the VP1 gene from a Chinese strain of
RA hepatitis A virus (HAV).";
RA Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14566; CAA32794.1; -
DR PIR; S04137; S04137.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1
FT CHAIN <1 1 COAT PROTEIN VP3 (1C).
FT CHAIN 2 340 COAT PROTEIN VP1 (1D).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;

Query Match 74.3%; Score 75; DB 1; Length 341;
Best Local Similarity 93.3%; Pred. No. 5.3e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 VLPPPRKMKGLFSQA 15
327 VLPPPRKMKGLFSQS 341

RESULT 11
DPYD_HUMAN
ID DPYD_HUMAN STANDARD; PRT; 1025 AA.
AC Q12882; Q16694; Q16761.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dihydropyrimidine dehydrogenase [NADP+] precursor (EC 1.3.1.2) (DPD)
DE (DHPDase) (Dihydrouracil dehydrogenase) (Dihydrothymine
DE dehydrogenase).
CN DPYD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94365020; PubMed=8083224;
```

```
RA Yokota H., Fernandez-Salguero P., Furuya H., Lin K., McBride O.W.,
RA Podschun B., Schnackerz K.D., Gonzalez F.J.;
RA "CDNA cloning and chromosome mapping of human dihydropyrimidine
RA dehydrogenase, an enzyme associated with 5-fluorouracil toxicity and
RA congenital thymine uraciluria.";
RA J. Biol. Chem. 269:23192-23196(1994).
RL [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97280676; PubMed=9135003;
RA Johnson M.R., Wang K., Tillmanns S., Albin N., Diasio R.B.;
RA "Structural organization of the human dihydropyrimidine dehydrogenase
RA gene.";
RA Cancer Res. 57:1660-1663(1997).
RL [3]
RP SEQUENCE OF 581-635 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97047101; PubMed=8892022;
RA Vreken P., van Kuilenburg A.B.P., Meisma R., Smit G.P.A.,
RA Bakker H.D., de Abreu R.A., van Gennip A.H.;
RA "A point mutation in an invariant splice donor site leads to exon
RA skipping in two unrelated Dutch patients with dihydropyrimidine
RA dehydrogenase deficiency.";
RA J. Inher. Metab. Dis. 19:645-654(1996).
RL [4]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92381021; PubMed=1512248;
RA Lu Z.-H., Zhang R., Diasio R.B.;
RA "Purification and characterization of dihydropyrimidine dehydrogenase
RA from human liver.";
RA J. Biol. Chem. 267:17102-17109(1992).
RL [5]
RP VARIANTS ARG-29: TRP-235 AND HIS-886.
RX MEDLINE=98102836; PubMed=9439663;
RA Vreken P., van Kuilenburg A.B.P., Meisma R., van Gennip A.H.;
RA "Dihydropyrimidine dehydrogenase (DPD) deficiency: Identification and
RA expression of missense mutations C29R, R886H and R235W.";
RA Hum. Genet. 101:333-338(1997).
RL [6]
RP VARIANTS ARG-29: TRP-235 AND HIS-886.
RX MEDLINE=97411371; PubMed=9266349;
RA Vreken P., van Kuilenburg A.B.P., Meisma R., van Gennip A.H.;
RA "Identification of novel point mutations in the dihydropyrimidine
RA dehydrogenase gene.";
RA J. Inher. Metab. Dis. 20:335-338(1997).
CC -1- FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE
CC REDUCTION OF URACIL AND THYMINE. ALSO INVOLVED THE DEGRADATION OF
CC THE CHEMOTHERAPEUTIC DRUG 5-FLUOROURACIL.
CC -1- CATALYTIC ACTIVITY: 5,6-dihydrouracil + NADP(+) = uracil + NADPH.
CC -1- COFACTOR: TWO EACH OF FAD AND FMN. ALSO CONTAINS TWO 4FE-4S
CC CLUSTERS. CONTAINS APPROXIMATELY 33 IRON ATOMS PER MOLECULE.
CC -1- PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF
CC URACIL AND THYMINE CATABOLISM AND IN THE PATHWAY LEADING TO THE
CC FORMATION OF BETA-ALANINE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES WITH GREATEST ACTIVITY
CC FOUND IN LIVER AND PERIPHERAL BLOOD MONONUCLEAR CELLS.
CC -1- DISEASE: DEFECTS IN DPYD ARE THE CAUSE OF HEREDITARY THYMINE-
CC URACILURIA (ALSO KNOWN AS FAMILIAL PYRIMIDINEMIA), A DISEASE
CC CHARACTERIZED BY PERSISTENT URINARY EXCRETION OF EXCESSIVE AMOUNTS
CC OF URACIL, THYMINE AND 5-HYDROXYMETHYLURACIL. PATIENTS SUFFERING
CC FROM THIS DISEASE SHOW A SEVERE REACTION TO THE ANTICANCER DRUG 5-
CC FLUOROURACIL. THIS REACTION INCLUDES STOMATITIS, LEUKOPENIA,
CC THROMBOCYTOPENIA, HAIR LOSS, DIARRHEA, FEVER, MARKED WEIGHT LOSS,
CC CEREBELLAR ATAXIA, AND NEUROLOGIC SYMPTOMS; PROGRESSING TO
CC SEMICOMA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

```
CC EMBL: U09178; AAA57474.1; -
CC EMBL: U20938; AAB51366.1; -
CC EMBL: X95670; CAA64973.1; -
CC GenBank: HGNC:3012; DPYD.
CC MIN: 274270; -
CC InterPro: IPR001450; 4Fe4S_ferredoxin.
CC InterPro: IPR001295; DHO_dh.
CC InterPro: IPR003009; FMN_enzyme.
CC Pfam: PF00037; fer4; 2.
CC Pfam: PF01180; DHodehase; 1.
CC ProDom: PD000139; FAD_pyr_redox; 1.
CC TIGRFAMs: TIGR01037; pyrd_subl_fam; 1.
CC PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
CC Oxidoreductase; NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4Fe-4S;
KW Disease mutation.
KW PROPEP 1 3
FT CHAIN 4 1025 DIHYDROPYRIMIDINE DEHYDROGENASE [NADP+].
FT NP_BIND 335 351 NADP (POTENTIAL).
FT NP_BIND 471 481 FAD (POTENTIAL).
FT DOMAIN 561 678 URACIL-BINDING (POTENTIAL).
FT METAL 953 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT VARIANT 29 29 C -> R (IN DPYD*9A AND DPYD*9B; LOSS OF ACTIVITY).
FT /FTID=VAR_005173.
FT R -> W (IN DPYD*8; LOSS OF ACTIVITY).
FT /FTID=VAR_005174.
FT S -> N (IN DPYD*4; LOW ACTIVITY).
FT /FTID=VAR_005175.
FT I -> V (IN DPYD*5).
FT /FTID=VAR_005176.
FT R -> H (IN DPYD*9B; 25% OF ACTIVITY).
FT /FTID=VAR_005177.
FT V -> F (IN DPYD*10; LOW ACTIVITY).
FT /FTID=VAR_005178.
FT S -> N (IN REF. 2).
FT CONFLICT 910 910
SQ SEQUENCE 1025 AA; 111374 MW; 521C9430C7F69AFA CRC64;
Query Match 52.0%; Score 52.5; DB 1; Length 1025;
Best Local Similarity 63.2%; Pred. No. 0.9;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPEKMEAYS-AKIALF 193
RESULT 12
DPYD_BOVIN STANDARD; PRT; 1025 AA.
AC Q28007; Q9TRV4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2) (DPD)
DE (DHPDHase) (Dihydrouracil dehydrogenase) (Dihydrothymine
DE dehydrogenase).
GN DPYD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97069948; PubMed=8912928;
RA Albin N., Johnson M.R., Diasio R.B.;
RT "CDNA cloning of bovine liver dihydropyrimidine dehydrogenase.";
RL DNA Seq. 6:243-250(1996).
RN [2]
RP SEQUENCE OF 668-678, AND ENZYME REGULATION.
RC TISSUE=Liver;
RX MEDLINE=92041818; PubMed=1939061;
RA Porter D.J.T., Chestnut W.G., Taylor L.C.E., Merrill B.M., Spector T.;
RT "Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.";
RL J. Biol. Chem. 266:19988-19994(1991).
CC -!- FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE
CC REDUCTION OF URACIL AND THYMINE.
CC -!- CATALYTIC ACTIVITY: 5,6-dihydrouracil + NADP(+) = uracil + NADPH.
CC -!- COFACTOR: FAD AND FMN. ALSO CONTAINS TWO 4FE-4S CLUSTERS.
CC -!- ENZYME REGULATION: INACTIVATED BY 5-IOUDOURACIL.
CC -!- PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF
CC URACIL AND THYMIDINE CATABOLISM AND IN THE PATHWAY LEADING TO THE
CC FORMATION OF BETA-ALANINE.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U20981; AAB40985.1; -
CC InterPro: IPR001450; 4Fe4S_ferredoxin.
CC InterPro: IPR001295; DHO_dh.
CC InterPro: IPR001327; FAD_pyr_redox.
CC InterPro: IPR003009; FMN_enzyme.
CC Pfam: PF00037; fer4; 2.
CC Pfam: PF01180; DHodehase; 1.
CC TIGRFAMs: TIGR01037; pyrd_subl_fam; 1.
CC ProDom: PD000139; FAD_pyr_redox; 1.
CC PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
CC Oxidoreductase; NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4Fe-4S.
FT METAL 953 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
SQ SEQUENCE 1025 AA; 111696 MW; 5B55F93A06C47F4F CRC64;
Query Match 45.0%; Score 45.5; DB 1; Length 1025;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 2 LPPPRKMKGLFSQAKISL 19
Db 176 LPPPEKMEAYS-AKIAL 192
RESULT 13
RLIX_ORYSA STANDARD; PRT; 178 AA.
AC Q943F3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ribosomal protein L18a.
GN RPL18A.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```


Query Match 43.6%; Score 44; DB 1; Length 340;
Best Local Similarity 61.5%; Pred. No. 7.2;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 PPRKMKGLFSQA 15
||| :: ||| ||
Db 242 PPRDRVYGLFROA 254

Search completed: February 3, 2003, 13:12:08
Job time : 6.16667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:07:47 : Search time 21.5 Seconds
(without alignments)
191.672 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	101	100.0	94 12 Q9ENW1	Q9enw1 hepatitis a
2	101	100.0	94 12 Q9ENW0	Q9enw0 hepatitis a
3	101	100.0	94 12 Q9ENV9	Q9env9 hepatitis a
4	101	100.0	94 12 Q9ENV8	Q9env8 hepatitis a
5	101	100.0	94 12 Q9ENV7	Q9env7 hepatitis a
6	101	100.0	94 12 Q9ENV6	Q9env6 hepatitis a
7	101	100.0	94 12 Q9ENV5	Q9env5 hepatitis a
8	101	100.0	94 12 Q9ENV4	Q9env4 hepatitis a
9	101	100.0	94 12 Q9ENV3	Q9env3 hepatitis a
10	101	100.0	94 12 Q9ENV2	Q9env2 hepatitis a
11	101	100.0	94 12 Q9ENV1	Q9env1 hepatitis a
12	101	100.0	94 12 Q9ENU9	Q9enu9 hepatitis a
13	101	100.0	94 12 Q9ENU8	Q9enu8 hepatitis a
14	101	100.0	94 12 Q9ENU7	Q9enu7 hepatitis a
15	101	100.0	94 12 Q9ENU6	Q9enu6 hepatitis a
16	101	100.0	94 12 Q9ENU5	Q9enu5 hepatitis a

17	101	100.0	94 12 Q9ENU4	Q9enu4 hepatitis a
18	101	100.0	94 12 Q9ENU2	Q9enu2 hepatitis a
19	101	100.0	94 12 Q9ENU1	Q9enu1 hepatitis a
20	101	100.0	94 12 Q9ENT9	Q9ent9 hepatitis a
21	101	100.0	94 12 Q9ENT5	Q9ent5 hepatitis a
22	101	100.0	94 12 Q9ENT1	Q9ent1 hepatitis a
23	101	100.0	115 12 Q9DWR4	Q9dwr4 hepatitis a
24	101	100.0	116 12 Q9W7X7	Q9w7x7 hepatitis a
25	101	100.0	116 12 Q71977	Q71977 hepatitis a
26	101	100.0	116 12 Q92941	Q92941 hepatitis a
27	101	100.0	116 12 Q9W7S7	Q9w7s7 hepatitis a
28	101	100.0	126 12 Q98W19	Q98w19 hepatitis a
29	101	100.0	132 12 Q8V4N6	Q8v4n6 hepatitis a
30	101	100.0	132 12 Q8V4N3	Q8v4n3 hepatitis a
31	101	100.0	132 12 Q8V4N2	Q8v4n2 hepatitis a
32	101	100.0	132 12 Q8V4N8	Q8v4n8 hepatitis a
33	101	100.0	132 12 Q8V4M7	Q8v4m7 hepatitis a
34	101	100.0	132 12 Q8V4M5	Q8v4m5 hepatitis a
35	101	100.0	132 12 Q8V4M4	Q8v4m4 hepatitis a
36	101	100.0	132 12 Q8V4M2	Q8v4m2 hepatitis a
37	101	100.0	132 12 Q8V4M1	Q8v4m1 hepatitis a
38	101	100.0	132 12 Q8V4L9	Q8v4l9 hepatitis a
39	101	100.0	132 12 Q8V4L8	Q8v4l8 hepatitis a
40	101	100.0	132 12 Q8V4L7	Q8v4l7 hepatitis a
41	101	100.0	132 12 Q8V4L6	Q8v4l6 hepatitis a
42	101	100.0	132 12 Q8V4L5	Q8v4l5 hepatitis a
43	101	100.0	132 12 Q8V4L3	Q8v4l3 hepatitis a
44	101	100.0	132 12 Q8V4L0	Q8v4l0 hepatitis a
45	101	100.0	132 12 Q8V4K8	Q8v4k8 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENW1 ID Q9ENW1 PRELIMINARY: PRT: 94 AA.
AC Q9ENW1;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus
OX NCBI_taxid=i2092;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN-NAGASAKI 01;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038276; BAB11814.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

|||||
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 2

Q9ENW0 ID Q9ENW0 PRELIMINARY: PRT: 94 AA.
AC Q9ENW0;
DT 01-MAR-2001 (Tremblrel. 16, Created)

```
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NAGASAKI 02;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038277; BAB11815.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10903 MW; 9AF8BE853D57329C CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 74 VLPPrKMKGLFSQAKISLF 93

RESULT 3
Q9ENV9 PRELIMINARY; PRT; 94 AA.
AC Q9ENV9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NAGASAKI 03;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038278; BAB11816.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFD4E8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 74 VLPPrKMKGLFSQAKISLF 93

RESULT 4
Q9ENV8 PRELIMINARY; PRT; 94 AA.
AC Q9ENV8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

```
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NAGASAKI 04;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038279; BAB11817.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 74 VLPPrKMKGLFSQAKISLF 93

RESULT 5
Q9ENV7 PRELIMINARY; PRT; 94 AA.
AC Q9ENV7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NAGASAKI 06;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038280; BAB11818.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 74 VLPPrKMKGLFSQAKISLF 93

RESULT 6
Q9ENV6 PRELIMINARY; PRT; 94 AA.
AC Q9ENV6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

RA Ida S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038281; BAB11819.1; -
FT NON_TER 1
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPrKMKGLFSQAKISLF 20
|||||
Db 74 VLPPrKMKGLFSQAKISLF 93

RESULT 7

Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGASAKI 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038282; BAB11820.1; -
FT NON_TER 1
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPrKMKGLFSQAKISLF 20
|||||
Db 74 VLPPrKMKGLFSQAKISLF 93

RESULT 8

Q9ENV4 PRELIMINARY; PRT; 94 AA.
AC Q9ENV4
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGASAKI 10;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038283; BAB11821.1; -

FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPrKMKGLFSQAKISLF 20
|||||
Db 74 VLPPrKMKGLFSQAKISLF 93

RESULT 9

Q9ENV3 PRELIMINARY; PRT; 94 AA.
AC Q9ENV3
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGASAKI 12N;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038284; BAB11822.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPrKMKGLFSQAKISLF 20
|||||
Db 74 VLPPrKMKGLFSQAKISLF 93

RESULT 10

Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NAGASAKI 15;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038285; BAB11823.1; -
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 11

Q9ENV1 ID Q9ENV1 PRELIMINARY; PRT; 94 AA.

AC Q9ENV1; 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGASAKI 18;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT Infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038286; BAB11824.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 12

Q9ENV9 ID Q9ENV9 PRELIMINARY; PRT; 94 AA.

AC Q9ENV9; 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGASAKI 21;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT Infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038288; BAB11826.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||

Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 13

Q9ENU8 ID Q9ENU8 PRELIMINARY; PRT; 94 AA.

AC Q9ENU8; 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGASAKI 22;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT Infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038289; BAB11827.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 14

Q9ENU7 ID Q9ENU7 PRELIMINARY; PRT; 94 AA.

AC Q9ENU7; 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGASAKI 23;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT Infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038290; BAB11828.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 15

Q9ENU6

ID Q9ENU6 PRELIMINARY; PRT; 94 AA.
AC Q9ENU6;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TremBLrel. 16, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAGASAKI 24;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038291; BAB11829.1; -.
BT NON_TER 1
BT NON_TER 94
BT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

Search completed: February 3, 2003, 13:14:22
Job time : 21.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 ; Search time 27.8333 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSOAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	AAW42926	Immunogenic Hepati
2	101	100.0	21	AAW42926	Synthetic HAV P2A
3	101	100.0	854	AAW50287	Sequence encoded b
4	101	100.0	1077	AAW95559	A partial hepatitis
5	101	100.0	1091	AAW32426	Translated from 5'
6	101	100.0	2227	AAW05697	Attenuated hepatitis
7	101	100.0	2227	AAW34074	Hepatitis A virus
8	101	100.0	2227	AAW18607	Amino acid sequenc
9	101	100.0	2227	AAW18608	Amino acid sequenc
10	101	100.0	2227	AAW18609	Amino acid sequenc

11	101	100.0	2227	23	AAE19899	Hepatitis A virus
12	98	97.0	2227	7	AAW60066	Sequence of viral
13	97	96.0	366	6	AAW50230	Sequence of Hepati
14	97	96.0	993	6	AAW50116	Sequence of Hepati
15	97	96.0	993	6	AAW50231	Sequence encoded b
16	49.5	49.0	20	19	AAW76539	Plasmid pGEM2-HAV-
17	47	46.5	518	20	AAW27340	Group B Streptococ
18	47	46.5	816	20	AAW27339	Group B Streptococ
19	47	46.5	846	23	ABP30554	Streptococcus poly
20	47	46.5	847	23	ABP27030	Streptococcus poly
21	47	46.5	847	23	ABP29773	Streptococcus poly
22	43	42.6	171	23	ABP40360	Staphylococcus epi
23	42	41.6	20	18	AAW42925	Immunogenic Hepati
24	42	41.6	21	22	AAW69442	Synthetic HAV P2A
25	42	41.6	125	23	ABW49486	Listeria monocytog
26	42	41.6	172	23	ABP25691	Streptococcus poly
27	42	41.6	503	22	ABG15762	Novel human diagno
28	41	40.6	144	22	AAW96238	Putative P. abyss
29	41	40.6	173	22	AAW49279	Protein encoded by
30	41	40.6	238	22	ABW68889	Drosophila melanog
31	41	40.6	2431	13	AAW251138	SFV4 non-structura
32	40	39.6	105	21	AAW19719	Arabidopsis thalia
33	40	39.6	105	21	AAW61014	Arabidopsis thalia
34	40	39.6	122	21	AAW41997	Human PRO polypt
35	40	39.6	173	22	AAW29276	Human PRO polypt
36	40	39.6	173	22	AAW49278	Protein encoded by
37	40	39.6	681	22	ABW63032	Drosophila melanog
38	40	39.6	748	22	AAW58340	Human brain expres
39	39.5	39.1	126	22	AAW58545	Propionibacterium
40	39	38.6	129	23	ABW90694	Chlamydia pneumoni
41	39	38.6	148	20	AAW35073	G. max CBF homolog
42	39	38.6	170	20	AAW33564	Drosophila melanog
43	39	38.6	200	22	ABW59162	Callithrix jacchus
44	39	38.6	216	22	AAW38461	Novel human diagno
45	39	38.6	222	22	ABW08469	

ALIGNMENTS

RESULT 1

AAW42926
ID AAW42926 standard; peptide; 20 AA.
XX
AC AAW42926;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1319.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Fields HA, Khudyakov YE;
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX PS Claim 18; Page 112; 140pp; English.

XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to

CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are

CC substantially similar to a portion of the amino acid sequence of the P2A

CC protein of HAV corresponding to amino acids 792-980. The present peptide

CC is derived from amino acids 834-853, and has a reactivity of 27.1% with

CC acute sera. Compositions containing the peptides can be used to induce an

CC immune response to HAV in a mammal. The peptides can also be used to

CC detect the presence of antibodies against HAV in mammalian serum. The

CC peptides can also be used to make an antibody against HAV by

CC administering the peptide to a mammal.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 101; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 6.7e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSQAKISLFYTEEHEIMKFS 20

1 FSQAKISLFYTEEHEIMKFS 20

Db

RESULT 2

AAB69443

ID AAB69443 standard; Peptide; 21 AA.

AC AAB69443;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 43.

DE Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;

KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX 25-JAN-2001.

XX 14-JUL-2000; 2000WO-US19267.

XX 15-JUL-1999; 99US-0144412.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

PI WPI; 2001-112681/12.

DR Synthetic peptides used as antigen sources for enzyme immunoassays

XX detecting anti-hepatitis A virus and as vaccines -

XX Claim 13; Page 95; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are

CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides

CC comprise antigenic epitopes of the major structural capsid polypeptides

CC or non-structural polypeptides of HAV with one or more glutamine

CC molecules at the carboxy end of the peptide. The peptides are used to

CC detect the presence of antibodies against HAV in mammalian serum, to

CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting

CC IgM antibodies in mammalian serum and detecting convalescence in a

CC mammal. The peptides are used to detect or quantify HAV antibodies in

CC samples in clinical or research-based assays using immunoblotting,

CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or biofluorescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy

CC end of the peptides enhances the IgM antibody reactivity.

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 101; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 7.1e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSQAKISLFYTEEHEIMKFS 20

1 FSQAKISLFYTEEHEIMKFS 20

Db

RESULT 3

AAP50287

ID AAP50287 standard; Protein; 854 AA.

XX AAP50287;

XX 30-NOV-1991 (first entry)

XX Sequence encoded by hepatitis A virus (HAV) cDNA from near the

DE genome 5' terminus to the end of the area corresponding to the

DE capsid protein region of poliovirus RNA.

XX Hepatitis A virus assay; antigen; antibody.

XX Hepatitis A virus.

OS WO8501517-A.

XX 11-APR-1985.

XX 27-SEP-1984; 84WO-US01552.

XX 30-SEP-1983; 83US-0537911.

XX (MASI) MASSACHUSETTS INST TECH.

XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;

PI Racanelliello VR;

XX WPI; 1985-098846/16.

DR N-PSDB; AAN50330.

XX New hepatitis A virus cDNA - useful in assays for the virus and

PT for prodn. of the viral antigen and antibodies to it

XX Example; Fig 7; 60pp; English.

XX The inventors claim HAV cDNA and a method for producing it, whereby

CC large amts. can be obtd. economically. The cDNA is useful in the

CC assay for detection of HAV quickly and easily and with high

CC sensitivity and specificity. The HAV cDNA is also used in the prodn.

CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.

XX SQ Sequence 854 AA;

Query Match 100.0%; Score 101; DB 6; Length 854;

Best Local Similarity 100.0%; Pred. No. 4.3e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSQAKISLFYTEEHEIMKFS 20

834 FSQAKISLFYTEEHEIMKFS 853

Db

RESULT 4

AAW95559

ID AAW95559 standard; Protein; 1077 AA.

XX

AAW95559;
28-APR-1999 (first entry)
A partial hepatitis A virus (HAV) protein.
Hepatitis A virus protein; HAV; P2 region;
cell-culture-adapted HAV strain; infection; accelerated growth.
Hepatitis A virus.
US5849562-A.
15-DEC-1998.
06-JUN-1995; 95US-0468926.
06-NOV-1991; 91US-0788262.
30-SEP-1983; 83US-0536911.
27-SEP-1984; 84US-0654942.
06-OCT-1988; 88US-0256135.
06-OCT-1988; 88US-0256135.
06-JUN-1995; 95US-0468926.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
Emerson SU, Purcell RH;
WPI: 1999-094412/08.
N-PSDB; AAX01006.
Chimeric hepatitis A virus strains - with P2 region from
cell-culture-adapted strain in wild-type genome
Disclosure; Fig 7A-L; 36pp; English.
The present sequence represents a partial hepatitis A virus (HAV)
protein. The specification describes a DNA construct consisting
of a wild-type HAV genome in which the P2 region is replaced by the
P2 region from a cell-culture-adapted HAV strain. The construct is
used to demonstrate that mutations in the P2 region of a
cell-culture-adapted HAV strain are sufficient for establishment of
infection and accelerated growth in cell culture.
SQ Sequence 1077 AA;
Query Match 100.0%; Score 101; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FSQAKISLFYTEEHEIMKFS 20
|||||
Db 1057 FSQAKISLFYTEEHEIMKFS 1076
RESULT 5
AAR32426
ID AAR32426 standard; Protein; 1091 AA.
XX
AC AAR32426;
XX
XX 17-DEC-2001 (updated)
DT 10-JUN-1993 (first entry)
XX
XX Translated from 5' region of Hepatitis A Virus genomic clone.
XX
KW HAV HM-175; chronic liver disease; picornavirus.
XX
OS Hepatitis A Virus.
XX
XX
FH Key Location/Qualifiers
FT Region 238..1091
FT /label= ORF
FT /note= "second putative initiation codon at

Region 1..711 position 240"
/note= "X's correspond to nonsense codons,
i.e. this region is not an ORF"
USN7788262-N.
15-DEC-1992.
30-SEP-1983; 83US-0536911.
27-SEP-1984; 84US-0654942.
06-OCT-1988; 88US-0256135.
30-SEP-1983; 83US-0536911.
06-NOV-1991; 91US-0788262.
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
PA
XX Baltimore D, Feinstein SM;
PI Purcell RH, Racaniello VR, Ticehurst JR;
XX
XX WPI: 1993-067429/08.
DR N-PSDB; AAQ36934.
XX
XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
of antigen and antibodies
Disclosure; Fig 7; 65pp; English.
XX
CC HAV virion RNA was extracted from the livers of marmosets which had
been inoculated with HAV (the HAV had previously been passaged twice
in marmosets). The RNA was used to prepare ds cDNA clones by
standard methods. Clones contg. inserts which hybridised to RNA from
HAV-infected African Green Monkey kidney cells were selected for
further analysis. A 7.4kb restriction map (about 99% of the HAV
genome) was constructed from 5 overlapping inserts. The sequence of
the first 3.3kb (approx.) from the 5'-terminus was determined. An
amino acid sequence was deduced from the entire clone and an open
reading frame was identified starting at position 238. A comparison
of the predicted HAV amino acid sequences with the known capsid
protein sequences of other picornaviruses (poliovirus, foot and
mouth disease virus and encephalomyelitis virus) revealed areas of
local homology.
CC (Note: Revised entry submitted to correct the patent number format of
US Government-owned NTIS applications to prevent clashes with ongoing US
granted patent numbers. For further information please visit the Derwent
web site at www.derwent.com/dwpl/updates/ntis_us.html.)
XX
SQ Sequence 1091 AA;
Query Match 100.0%; Score 101; DB 14; Length 1091;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FSQAKISLFYTEEHEIMKFS 20
|||||
Db 1071 FSQAKISLFYTEEHEIMKFS 1090
RESULT 6
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX
AC AAR05697;
XX
XX 15-AUG-1990 (first entry)
DT
XX
XX Attenuated hepatitis A virus.
XX
XX
KW Hepatitis A virus; vaccine; attenuated.
XX
OS Hepatitis A virus, strain HM-175.
XX

```

FH Key Location/Qualifiers
FT Region 1..23
FT /label=VP4 = 1A
FT Region 24..245
FT /label=VP2 = 1B
FT Region 246..491
FT /label=VP3 = 1C
FT Region 492..791
FT /label=VP1 = 1D
FT Region 792..980
FT /label=2A
FT Region 981..1087
FT /label=2B
FT Region 1088..1422
FT /label=2C
FT Region 1423..1496
FT /label=3A
FT Region 1497..1519
FT /label=3B = VPg
FT Region 1520..1738
FT /label=3C
FT Region 1739..2227
FT /label=3D
XX US4894228-A.
XX PN
XX 16-JAN-1990.
XX PD
XX 12-JUL-1988; 88US-0217824.
XX PR
XX 12-JUL-1988; 88US-0217824.
XX PR
XX 12-JUL-1988; 88US-0652967.
XX PA
XX (USSH ) US DEPT HEALTH & HUMAN.
XX XX
XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstone SM;
XX Daemer RJ, Gust ID;
XX WPI: 1990-075557/10.
XX DR
XX N-PSDB; AAQ03512.
XX XX
XX Vaccine against hepatitis A virus infection - comprises novel
XX attenuated hepatitis A virus strain.
XX Claim 1; Fig 1; 18pp; English.
XX XX
XX The attenuated HAV is useful for inducing protective immunity against
XX HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
XX several nucleotide changes distributed throughout the genome, is
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is
XX suitable for use as an HAV vaccine. It is noted that not all the changes
XX are necessary for attenuation and use as a vaccine.
XX SQ
XX Sequence 2227 AA;
XX Query Match 100.0%; Score 101; DB 11; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSQAKISLFYTEHEIMKFS 20
DB 834 FSQAKISLFYTEHEIMKFS 853
RESULT 7
AAW34074
ID AAW34074 standard; Protein; 2227 AA.
XX AC
XX AAW34074;
XX XX
XX 27-APR-1998 (first entry)
XX XX
XX Hepatitis A virus HM-175 protein sequence.

```

```

XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
XX infection; vaccine.
XX OS
XX Hepatitis A virus HM-175.
XX Key Location/Qualifiers
XX 1..23
XX FT Protein /label= VP4
XX FT Protein 24..245
XX FT /label= VP2
XX FT Protein 246..491
XX FT /label= VP3
XX FT Protein 492..791
XX FT /label= VP1
XX FT Protein 792..980
XX FT /label= 2A
XX FT Protein 981..1087
XX FT /label= 2B
XX FT Protein 1088..1422
XX FT /label= 2C
XX FT Protein 1423..1496
XX FT /label= 3A
XX FT Protein 1497..1519
XX FT /label= 3B
XX FT Protein 1520..1738
XX FT /label= 3C
XX FT Protein 1739..2227
XX FT /label= 3D
XX XX
XX WO9740166-A2.
XX PN
XX 30-OCT-1997.
XX PD
XX 18-APR-1997; 97WO-US06506.
XX PF
XX 19-APR-1996; 96US-0015642.
XX PR
XX (USSH ) US SEC DEPT HEALTH.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX XX
XX Emerson SU, Purcell RH, Raychaudhuri G;
XX WPI: 1997-535850/49.
XX DR
XX N-PSDB; AAT93023.
XX XX
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
XX as vaccines against HAV infection
XX Disclosure; Fig 13A-D; 66pp; English.
XX XX
XX This protein sequence is encoded by the human hepatitis A virus
XX (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
XX HAV/7 is obtained by passage of HM-175 in African Green Monkey
XX kidney cells. A claimed DNA construct (I) comprises a genome of
XX HAV, where the genome is a human attenuated HAV genome in which a
XX region of the 2C gene has been replaced by a corresponding region
XX from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
XX region of the 2C gene from AGM-27 contained in the construct
XX preferably encodes amino acids 120-328 of the 2C protein, amino
XX acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
XX transcript of (1); (2) a cell transfected with (1) or the RNA
XX transcript of (1); (3) a HAV genome as above; (4) antibodies to the
XX HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
XX its RNA transcript, can be used as a vaccine for preventing HAV in
XX a mammal. (I) or the RNA transcript can also be used to stimulate
XX the production of protective antibodies in the mammal.
XX SQ
XX Sequence 2227 AA;
XX Query Match 100.0%; Score 101; DB 18; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FSQAKISLFYTEEHEIMKFS 20
   |||||
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 8
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
XX AC AAB18607;
XX DT 15-JAN-2001 (first entry)
XX DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX OS Hepatitis A virus.
   |||||
XX PN US6113912-A.
XX PD 05-SEP-2000.
XX PF 07-JUN-1995; 95US-0475886.
XX PR 18-SEP-1992; 92US-0947338.
XX PR 17-SEP-1993; 93WO-US08610.
XX PR 10-MAR-1995; 95US-0397232.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX DR WPI; 2000-586464/55.
XX DR N-PSDB; AAA75476.
XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
   line useful as vaccine for protecting humans against hepatitis A virus
   infection, has modified genome compared to wild type -
XX PS Disclosure; Columns 67-78; 72pp; English.
XX CC The present sequence is derived from passage 35 of a wild type
   hepatitis A virus (HAV) strain HM-174. The resulting virus is
   designated P-35 virus. The sequence is modified to produce HAV which
   are adapted to growth in the human fibroblast-like cell line MRC-5.
   The HAV is able to propagate in MRC-5 cells and retain appropriate
   attenuation. It is useful as a live vaccine for prophylaxis of
   hepatitis A in humans and other primates.
XX PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX DR WPI; 2000-586464/55.
XX DR N-PSDB; AAA75476.
XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
   line useful as vaccine for protecting humans against hepatitis A virus
   infection, has modified genome compared to wild type -
XX PS Disclosure; Fig 6A-K; 72pp; English.
XX CC The present sequence is derived from a wild type hepatitis A virus
   (HAV) strain HM-174. The sequence is modified to produce HAV which
   are adapted to growth in the human fibroblast-like cell line MRC-5.
   The HAV is able to propagate in MRC-5 cells and retain appropriate
   attenuation. It is useful as a live vaccine for prophylaxis of
   hepatitis A in humans and other primates.
XX SQ Sequence 2227 AA;
   Query Match 100.0%; Score 101; DB 21; Length 2227;
   Best Local Similarity 100.0%; Pred. No. 1.3e-07;
   Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHEIMKFS 20
   |||||
Db 834 FSQAKISLFYTEEHEIMKFS 853

RESULT 9
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
XX AC AAB18608;
XX DT 15-JAN-2001 (first entry)
XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
   P-35 virus.

```

XX WPI: 2000-586464/55.
DR N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
XX Disclosure; Columns 93-104; 72pp; English.
XX
XX The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853
RESULT 11
AAE19899
ID AAE19899 standard; Protein; 2227 AA.
AC AAE19899;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis A virus (HAV) protein.
XX
KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
OS Hepatitis A virus.
XX
PN WO200213855-A2.
XX
PD 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB01808.
XX
PR 17-AUG-2000; 2000US-225767P.
PR 29-AUG-2000; 2000US-229175P.
PR 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
PI
DR WPI: 2002-241837/29.
DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus -
XX
XX Claim 11; Page 82-87; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to
CC a hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is

CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853
RESULT 12
AAP60066
ID AAP60066 standard; Protein; 2227 AA.
XX
AC AAP60066;
XX
XX 26-JUN-1991 (first entry)
DT
DE Sequence of viral L434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.
XX
KW Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
FH Key Location/Qualifiers
FT Region 1..245
FT /label= P1.1A
FT Region 246..491
FT /label= 1B
FT Region 492..836
FT /label= 1C
FT Region 837..980
FT /label= P2.2A
FT Region 981..1076
FT /label= 2B
FT Region 1077..1422
FT /label= 2C
FT Region 1423..1484
FT /label= P3.3A
FT Region 1485..1507
FT /label= 3B
FT Region 1508..1678
FT /label= 3C
FT Region 1679..2227
FT /label= 3D
XX
XX EP199480-A.
PN
XX 29-OCT-1986.
XX
XX 03-APR-1986; 86EP-0302465.
PF
XX
PR 03-APR-1985; 85US-0719329.
XX
XX (CHIR-) CHIRON CORP.
PA
XX
PI Dina D, Potter SJ, Vannest GA, Caput D;
XX
DR WPI: 1986-286213/44.
DR N-PSDB; AAN60080.
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use
PT in prodn. of vaccines and diagnostic probes
PS
XX Claim 5; Fig 1; 18pp; English.
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
CC hepatitis A virus; transformed hosts may be used for expression of

CC polypeptides and fragments useful in vaccines without risk of
 CC infection by the virus or in prodn. of particles which are capable
 CC of inducing immunocompetent B cells for passive immunotherapy. Pref.
 CC epitope is derived from AAs 445-657 or 792-848 of the HAV
 CC polypeptide sequence (AAP60066).

XX Sequence 2227 AA;

Query Match 97.0%; Score 98; DB 7; Length 2227;
 Best Local Similarity 95.0%; Pred. No. 3.9e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHEIMKFS 20
 |||||:|||||
 Db 834 FSQAKISLFYTDHEIMKFS 853

RESULT 13
 AAP50230
 ID AAP50230 standard; Protein; 366 AA.

XX AAP50230;

XX 28-NOV-1991 (first entry)

XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).

XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.

XX Hepatitis A virus.

XX EPI138704-A.

XX 24-APR-1985.

XX 09-OCT-1984; 84EP-0402025.

XX 02-MAR-1984; 84US-0585942.

XX 14-OCT-1983; 83US-0541836.

XX (MERI) MERCK & CO INC.

XX Hughes JV, Scolnick EM, Tomassini JE;

XX WPI; 1985-100818/17.

XX N-PSDB; AAN50274.

XX New hepatitis A virus surface protein - useful for binding to
 XX neutralising antibodies to the virus

XX Claim 21; Page 46-48; 49pp; English.

XX VP1 is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.

XX Sequence 366 AA;

Query Match 96.0%; Score 97; DB 6; Length 366;
 Best Local Similarity 100.0%; Pred. No. 7.7e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHEIMKFS 19
 |||||:|||||
 Db 343 FSQAKISLFYTEEHEIMKFS 361

RESULT 14
 AAP50116
 ID AAP50116 standard; Protein; 993 AA.

XX AAP50116;

XX 30-SEP-1991 (first entry)
 XX Sequence of Hepatitis A virus (HAV) immunogenic peptides
 DE VP-1, VP-2, VP-3 and VP-4.

XX Antigenic protein; immunogen; vaccine.

XX Hepatitis A virus (strain CR326).

XX EPI154587-A.

XX 11-SEP-1985.

XX 27-FEB-1985; 85EP-0400369.

XX 02-MAR-1984; 84US-0585818.

XX (MERI) MERCK & CO INC.

XX Linemeyer DL, Menke JG, Reuben RG, Mitra SW;

XX WPI; 1985-224964/37.

XX N-PSDB; AAN50139.

XX New nucleotide sequences coding for hepatitis A virus antigens -
 XX useful for eliciting normal immune response and in vaccines for
 XX protecting against the virus

XX Example; Page 11-17; 32pp; English.

XX Within the sequence in AAN50139 is encoded the information necessary
 CC to make the antigenic proteins of HAV. The sequences encoding for
 CC the structural proteins begin at base 403. The key sub-unit
 CC sequences within VP-1, designated Sequences I,II,III,IV, and V,
 CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
 CC nucleotide sequences which are valuable as encoding antigenic
 CC proteins are the sequences from base 1749 to base 2722; from base
 CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
 CC base 1749 to base 2722 is esp. valuable as a vector for producing
 CC antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
 CC translation of a stop codon.

XX Sequence 993 AA;

Query Match 96.0%; Score 97; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHEIMKFS 19
 |||||:|||||
 Db 970 FSQAKISLFYTEEHEIMKFS 988

RESULT 15
 AAP50231
 ID AAP50231 standard; Protein; 993 AA.

XX AAP50231;

XX 28-NOV-1991 (first entry)

XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).

XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.

XX Hepatitis A virus.

XX Key Location/Qualifiers

XX Protein 628..993

XX /note= "claimed; X denotes translated stop codons

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 ; Search time 9.5 Seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	1091	6	Patent No. 5516630
2	101	100.0	2227	3	US-08-475-886-2
3	101	100.0	2227	3	US-08-475-886-4
4	101	100.0	2227	3	US-08-475-886-6
5	101	100.0	2227	4	US-08-397-232-2
6	101	100.0	2227	4	US-08-397-232-4
7	101	100.0	2227	4	US-09-171-387-2
8	101	100.0	2227	4	US-09-653-499-2
9	101	100.0	2227	4	US-09-653-499-4
10	101	100.0	2227	4	US-09-653-499-6
11	43	42.6	171	4	US-09-134-001C-5205
12	41	40.6	2431	1	US-07-920-281C-2
13	41	40.6	2431	4	US-08-466-277-2
14	39	38.6	170	4	US-09-198-119C-89
15	39	38.6	289	5	PCT-US92-00282-11
16	39	38.6	860	1	US-08-092-817-4
17	39	38.6	860	4	US-08-485-128-4
18	39	38.6	860	4	US-09-804-778A-8
19	39	38.6	1410	2	US-08-470-058-4
20	39	38.6	1410	3	US-09-037-188-4
21	39	38.6	1410	4	US-09-285-310-4
22	38	37.6	265	1	US-07-958-551-2
23	38	37.6	265	1	US-08-129-610-7
24	38	37.6	265	1	US-08-129-609A-7
25	38	37.6	265	1	US-08-455-313-7
26	38	37.6	265	1	US-08-475-924-2
27	38	37.6	265	2	US-08-657-579A-2

```

28 38 37.6 265 4 US-09-224-025-7 Sequence 7, Appli
29 38 37.6 265 5 PCT-US94-07887-7 Sequence 7, Appli
30 38 37.6 707 3 US-08-772-270A-4 Sequence 4, Appli
31 37 36.6 189 4 US-09-134-001C-3160 Sequence 3160, Ap
32 37 36.6 198 1 US-08-097-827-2 Sequence 2, Appli
33 37 36.6 198 1 US-08-494-574-2 Sequence 2, Appli
34 37 36.6 420 3 US-08-685-871-58 Sequence 58, Appli
35 37 36.6 711 3 US-08-772-270A-12 Sequence 12, Appli
36 37 36.6 1354 3 US-08-685-871-2 Sequence 2, Appli
37 37 36.6 1388 2 US-08-685-576-1 Sequence 1, Appli
38 37 36.6 1388 2 US-08-685-576-4 Sequence 4, Appli
39 36 35.6 147 3 US-08-946-329A-62 Sequence 62, Appli
40 36 35.6 169 2 US-08-912-227-2 Sequence 2, Appli
41 36 35.6 169 4 US-09-513-584-2 Sequence 2, Appli
42 36 35.6 300 4 US-09-261-599B-6 Sequence 6, Appli
43 36 35.6 361 4 US-09-261-599B-1 Sequence 1, Appli
44 36 35.6 361 4 US-09-456-455A-1 Sequence 1, Appli
45 36 35.6 416 3 US-08-946-329A-17 Sequence 17, Appli

```

ALIGNMENTS

```

RESULT 1
5516630-2
: Patent No. 5516630
: APPLICANT: RICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
: STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
: BAROUDY, BAHIGE M.
: TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
: NUMBER OF SEQUENCES: 4
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/788,262
: FILING DATE: 06-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 256,135
: FILING DATE: 06-OCT-1988
: APPLICATION NUMBER: 654,942
: FILING DATE: 27-SEP-1984
: APPLICATION NUMBER: 537,911
: FILING DATE: 30-SEP-1983
: SEQ ID NO: 2
: LENGTH: 1091
5516630-2

```

```

Query Match 100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FSQAKISLFYTEHEIMKFS 20
|||||
Db 1071 FSQAKISLFYTEHEIMKFS 1090
|||||

```

```

RESULT 2
US-08-475-886-2
: Sequence 2, Application US/08475886A
: Patent No. 6113912
: GENERAL INFORMATION:
: APPLICANT: FUNKHOUSER, ANN W
: APPLICANT: EMERSON, SUZANNE U
: APPLICANT: PURCELL, ROBERT H
: APPLICANT: D'HONDT, ERIC
: TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
: FILE REFERENCE: 20264262US2
: CURRENT APPLICATION NUMBER: US/08/475,886A
: CURRENT FILING DATE: 1995-06-07
: EARLIER APPLICATION NUMBER: 07/947,338
: EARLIER FILING DATE: 1992-09-18
: EARLIER APPLICATION NUMBER: 08/397,232
: EARLIER FILING DATE: 1995-03-10
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent In Ver. 2.1

```

; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEEHEIMKFS 20
Db 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912

GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEEHEIMKFS 20
Db 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 4
US-08-475-886-6
; Sequence 5, Application US/08475886A
; Patent No. 6113912

GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEEHEIMKFS 20
Db 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSOAKISLFYTEEHEIMKFS 20
Db 834 FSOAKISLFYTEEHEIMKFS 853

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110

GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0

Qy      1 FSOAKISLFYTEEHEIMKFS 20
        |||||
Db      834 FSOAKISLFYTEEHEIMKFS 853

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499

```

```

Query Match      100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFYTEHEIMKFS 20
        |||||
Db      834 FSOAKISLFYTEHEIMKFS 853

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:14:27 ; Search time 6 Seconds
(without alignments)
67.262 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSOAKISLFYTEEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	9	US-10-104-966-12
2	101	100.0	2227	9	US-10-135-988-2
3	101	100.0	2227	9	US-10-135-988-4
4	101	100.0	2227	9	US-10-135-988-6
5	101	100.0	2227	10	US-09-329-955-12
6	41	40.6	2431	10	US-09-901-106-2
7	40	39.6	173	9	US-10-174-590-506
8	40	39.6	173	9	US-10-176-758-506
9	40	39.6	173	9	US-10-175-737-506
10	40	39.6	173	12	US-10-052-586-506
11	40	39.6	748	10	US-09-864-761-43244
12	39	38.6	170	9	US-09-996-140-89
13	39	38.6	216	10	US-09-747-155-92
14	39	38.6	289	10	US-09-305-8568-8
15	39	38.6	326	10	US-09-894-991-2
16	39	38.6	860	10	US-09-824-637-4
17	38	37.6	421	9	US-09-738-626-5589
18	38	37.6	456	10	US-09-815-242-11304
19	37	36.6	182	10	US-09-864-761-42913

Sequence 198, App
Sequence 5, Appli
Sequence 4, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 93, Appli
Sequence 47670, A
Sequence 2, Appli
Sequence 2, Appli
Sequence 16, Appli
Sequence 10234, A
Sequence 11720, A
Sequence 2, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 1, Appli
Sequence 2, Appli
Sequence 334, App
Sequence 86, Appl
Sequence 5158, Ap

20 37 36.6 432 10 US-09-881-752A-198
21 37 36.6 700 9 US-10-013-310-5
22 37 36.6 765 10 US-09-804-472-4
23 37 36.6 767 10 US-09-804-472-5
24 37 36.6 791 10 US-09-804-472-2
25 36 35.6 140 9 US-09-791-932-93
26 36 35.6 166 10 US-09-864-761-47670
27 36 35.6 169 9 US-10-136-511-2
28 36 35.6 169 10 US-09-345-790-2
29 36 35.6 177 12 US-10-080-455-1
30 36 35.6 177 12 US-10-116-378-16
31 36 35.6 244 10 US-09-815-242-10234
32 36 35.6 248 10 US-09-815-242-11720
33 36 35.6 300 9 US-10-077-698-6
34 36 35.6 361 9 US-10-086-181-2
35 36 35.6 361 9 US-09-995-225-8
36 36 35.6 361 9 US-10-077-698-1
37 36 35.6 361 12 US-10-015-498-2
38 36 35.6 406 12 US-10-062-254-332
39 36 35.6 416 8 US-08-900-220C-18
40 36 35.6 416 10 US-09-151-999-18
41 36 35.6 577 10 US-09-801-368-420
42 36 35.6 608 10 US-09-740-046-2
43 36 35.6 613 12 US-10-062-254-334
44 36 35.6 2026 10 US-09-801-368-86
45 35.5 35.1 129 10 US-09-815-242-5158

ALIGNMENTS

RESULT 1
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 FSOAKISLFYTEEHEIMKFS 20
Db 834 FSOAKISLFYTEEHEIMKFS 853
RESULT 2
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

```
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FSOAKISLFYTEHEIMKFS 20
|||||
DB 834 FSOAKISLFYTEHEIMKFS 853
```

```
RESULT 3
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FSOAKISLFYTEHEIMKFS 20
|||||
DB 834 FSOAKISLFYTEHEIMKFS 853
```

```
RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
```

```
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
```

```
Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FSOAKISLFYTEHEIMKFS 20
|||||
DB 834 FSOAKISLFYTEHEIMKFS 853
```

```
RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
```

```
Query Match 100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FSOAKISLFYTEHEIMKFS 20
|||||
DB 834 FSOAKISLFYTEHEIMKFS 853
```

```
RESULT 6
US-09-901-106-2
; Sequence 2, Application US/09901106
; Patent No. US20020151067A1
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; Lilljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
```

```

STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/901,106
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/920,281C
FILING DATE: 13-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-901-106-2

```

Query Match	40.6%	Score 41	DB 10	Length 2431
Best Local Similarity	50.0%	Pred. No. 1.2e+02		
Matches 7: Conservative	2	Mismatches 5	Indels 0	Gaps 0

```
QY      1 FSQAKISLFYTEEH 14
      || |::| |
Db      1072 FSAPKVSLYYENN 1085
```

```

RESULT 7
US-10-174-590-506
; Sequence 506, Application US/10174590
; Publication No. US2003008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-506

```

Query Match	39.6%	Score 40;	DB 9;	Length 173;
Best Local Similarity	35.0%	Pred. No. 9.7;		
Matches	7;	Conservative	7;	Mismatches
			6;	Indels
			0;	Gaps
			0;	

```
Qy 1 FSQAKISLFYTEEHEIMKFS 20
    ::::|::|
Db 148 WAQSLVSLFOALRHDLMRSS 167
```

```

RESULT 8
US-10-176-758-506
; Sequence 506, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-506

```

Query Match 39.6%; Score 40; DB 9; Length 173;
Best Local Similarity 35.0%; Pred. No. 9.7;
Matches 7; Conservative 7; Mismatches 6; Indels

```
QY 1 FSQAKISLFYTEEHEIMKFS 20
    ::::| | | | |:::|
Db 148 WAQSLVSLFQALRHDLMRSS 167
```

```

RESULT 9
US-10-175-737-506
; Sequence 506, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC50
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-506

```

Query Match 39.6%; Score 40; DB 9; Length 173;
Best Local Similarity 35.0%; Pred. No. 9.7;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLTYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLMRSS 167

RESULT 10

US-10-052-586-506
; Sequence 506, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22

;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655
;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 39.6%; Score 40; DB 12; Length 173;
Best Local Similarity 35.0%; Pred. No. 9.7;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHEIMKFS 20
::: ||| |::: |
Db 148 WAQSLVSLFQALRHDLMRSS 167

RESULT 11

US-09-864-761-43244

; Sequence 43244, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006657
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006654
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006659
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006655
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006658
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 43244
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC00486.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EST_HUMAN HIT: A1636014.1, EVALUATE 3.00e-51
; OTHER INFORMATION: SWISSPROT HIT: P38254, EVALUATE 1.40e+00
US-09-864-761-43244

Query Match 39.6%; Score 40; DB 10; Length 748;
Best Local Similarity 53.3%; Pred. No. 49;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 SOAKISLFYTEEHEI 16
||| |::: |
Db 346 SOKKTSLFDEPEL 360

```
RESULT 12
US-09-996-140-89
; Sequence 89, Application US/09996140
; Patent No. US20020157136A1
; GENERAL INFORMATION:
; APPLICANT: Thomashow, Michael
; APPLICANT: Stockinger, Eric
; APPLICANT: Jaglo-Ottosen, Kirsten
; APPLICANT: Gilmour, Sarah
; APPLICANT: Zarka, Daniel
; APPLICANT: Jiang, Cai-Zhong
; TITLE OF INVENTION: PLANT HAVING ALTERED ENVIRONMENTAL
; TITLE OF INVENTION: STRESS TOLERANCE
; FILE REFERENCE: MB10040US
; CURRENT APPLICATION NUMBER: US/09/996,140
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/706,270
; PRIOR FILING DATE: 1996-09-04
; PRIOR APPLICATION NUMBER: 09/018,233
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,816
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,235
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/017,575
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,227
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/018,234
; PRIOR FILING DATE: 1998-02-03
; PRIOR APPLICATION NUMBER: 09/198,119
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: PCT/US99/01895
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Glycine max
US-09-996-140-89

Query Match 38.6%; Score 39; DB 9; Length 170;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 SOAKISLFYTEEHE 15
:: | : |||| |
113 TEQKRGMYTEEEE 126

RESULT 13
US-09-747-155-92
; Sequence 92, Application US/09747155
; Patent No. US20020151692A1
; GENERAL INFORMATION:
; APPLICANT: Rouquier, Sylvie
; APPLICANT: Giorgi, Dominique
; TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding Sam
; FILE REFERENCE: 19904-008 (C009B6834US)
; CURRENT APPLICATION NUMBER: US/09/747,155
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/171,746
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 431
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Callithrix jacchus
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: (1)..(649)
; OTHER INFORMATION: Taxon = 9483; gene = CJA21; Accession DDBJ/EMBL/GenBank = AF12
US-09-747-155-92

Query Match 38.6%; Score 39; DB 10; Length 216;
Best Local Similarity 43.8%; Pred. No. 18;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 KISLFYTEEHEIMKFS 20
|| | : | : | : | : |
106 KIPQFFCEIHQIQFA 121

RESULT 14
US-09-305-856B-8
; Sequence 8, Application US/09305856B
; Patent No. US20020061518A1
; GENERAL INFORMATION:
; APPLICANT: Penny, Laura
; APPLICANT: Galvin, Margaret
; TITLE OF INVENTION: Genotyping the Human
; TITLE OF INVENTION: UDP-Glucuronosyltransferase 1 (UGT1) Gene
; FILE REFERENCE: 4389-7 (formerly SEQ-17C1P)
; CURRENT APPLICATION NUMBER: US/09/305,856B
; CURRENT FILING DATE: 1999-05-05
; PRIOR APPLICATION NUMBER: 60/084,807
; PRIOR FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-305-856B-8

Query Match 38.6%; Score 39; DB 10; Length 289;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 FYTEEHEIMKFS 20
| : | : | : | : |
100 FFETEHLLMKFS 111

RESULT 15
US-09-894-991-2
; Sequence 2, Application US/09894991
; Patent No. US2002090619A1
; GENERAL INFORMATION:
; APPLICANT: Pfeiffer, Douglas R.
; APPLICANT: Murphy, Anne M.
; APPLICANT: Jung, Dennis W.
; APPLICANT: Bradshaw, Patrick C.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING MITOCHONDRIAL
; TITLE OF INVENTION: DIVALENT CATION TRANSPORTERS
; FILE REFERENCE: 660088.430
; CURRENT APPLICATION NUMBER: US/09/894,991
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Acinetobacter sp.
US-09-894-991-2

Query Match 38.6%; Score 39; DB 10; Length 326;
Best Local Similarity 56.2%; Pred. No. 29;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEEHEI 16
| : | : | : | : |
276 FEKLKSSLVNTEYHEI 291
```

Search completed: February 3, 2003, 13:32:26
Job time : 7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 ; Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEEHIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1358	2 A03905	genome polyprotein
2	101	100.0	2227	1 GNNYHM	genome polyprotein
3	101	100.0	2227	1 GNNYHR	genome polyprotein
4	101	100.0	2227	1 GNNYHB	genome polyprotein
5	101	100.0	2227	1 GNNYHB	genome polyprotein
6	97	96.0	852	1 GNNYHA	genome polyprotein
7	88	87.1	2230	1 GNNYSA	genome polyprotein
8	47	46.5	378	2 F54300	formate dehydrogen
9	43	42.6	325	2 H71869	hypothetical prote
10	43	42.6	410	2 T22413	hypothetical prote
11	42	41.6	125	2 AB1747	glycine cleavage s
12	42	41.6	125	2 A11377	glycine cleavage s
13	42	41.6	187	2 T18844	hypothetical prote
14	42	41.6	545	2 S10842	gene ND5 intron 2
15	42	41.6	1265	2 F88886	protein F52G2.2 [1
16	41.5	41.1	288	2 S64243	hypothetical prote
17	41	40.6	144	2 E75128	glycine cleavage s
18	41	40.6	190	2 C90529	nuclease, lipoprot
19	41	40.6	209	2 S65300	hypothetical prote
20	41	40.6	304	2 C72324	homoserine O-succi
21	41	40.6	465	2 B97235	deacetylase/dipep
22	41	40.6	483	2 G84113	hypothetical prote
23	41	40.6	557	2 S58164	glucose-6-phosphat
24	41	40.6	820	2 T51510	hypothetical prote
25	41	40.6	1040	1 A38306	alpha-mannosidase
26	41	40.6	2431	1 MNVVSF	nonstructural poly
27	40.5	40.1	251	2 T13104	minor tail protein
28	40.5	40.1	419	2 T25577	hypothetical prote
29	40	39.6	142	2 C70059	hypothetical prote

ALIGNMENTS

RESULT 1

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Malzel Jr., J.V.; Purcell, R.H.; Feins

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 101; DB 2; Length 1358;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEEHIMKFS 20

Db 834 FSQAKISLFYTEEHIMKFS 853

RESULT 2

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core

B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 3
GNNYHR

genome polypeptide - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core protein 2B; coat protein 2C; coat protein 3A; coat protein 3B; coat protein 3C; coat protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest, P.; Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; PMID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAU>

C:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597

C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 4
GNNYMK

genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core protein 2B; coat protein 2C; coat protein 3A; coat protein 3B; coat protein 3C; coat protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.H.; Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with other HAV sequences.
A:Reference number: A94149; PMID:87175701; PMID:3031686
A:Accession: A94149
A>Status: nucleic acid sequence not shown

A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to Genbank, August 1987
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 5
GNNYHB

genome polypeptide - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; coat protein 2B; coat protein 2C; coat protein 3A; coat protein 3B; coat protein 3C; coat protein 3D
Vpg: protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, T.; Wimmer, E.; Deinhard, G.; Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso 1000)
A:Reference number: J50303; PMID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>

C:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 1D #status predicted <VP1>
F:981-1108/Product: core protein 2A #status predicted <P2A>
F:1109-1438/Product: core protein 2B #status predicted <P2B>
F:1439-1496/Product: core protein 2C #status predicted <P2C>
F:1497-1519/Product: protein 3A #status predicted <P3A>
F:1520-1736/Product: genome-linked protein vpg #status predicted <VPG>
F:1737-2227/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 6
GNNYHA

genome polypeptide - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03904
 R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.
 J. Virol. 54, 247-255, 1985
 A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
 A:Reference number: A03904; MUID:85185648; PMID:2985793
 A:Accession: A03904
 A:Molecule type: genomic RNA
 A:Residues: 1-852 <LIN>
 A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; coat protein 1A #status predicted <C1A>
 F:1-245/Product: coat protein 1A #status predicted <C1A>
 F:246-491/Product: coat protein 1B #status predicted <C1B>
 F:492-836/Product: coat protein 1C #status predicted <C1C>
 F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 96.0%; Score 97; DB 1; Length 852;
 Best Local Similarity 100.0%; Pred. No. 4.7e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 FSQAKISLFYTEHEIMKFS 19
 |||||
 Db 834 FSQAKISLFYTEHEIMKFS 852

RESULT 7
 GNNVSA
 genome polyprotein - simian hepatitis A virus (strain AGM-27)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
 C:Species: simian hepatitis A virus
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
 C:Accession: A30470; A04885; S03965
 R:Tzarev, S.A.
 Submitted to JIPID, April 1991
 A:Reference number: A30470
 A:Accession: A30470
 A:Molecule type: genomic RNA
 A:Residues: 1-2230 <TSA>
 A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
 R:Tzarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
 J. Gen. Virol. 72, 1677-1683, 1991
 A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
 A:Reference number: J01080; MUID:91311420; PMID:1649901
 A:Contents: annotation
 A:Note: neither amino acid nor nucleotide sequence is given
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
 submitted to the EMBL Data Library, May 1989
 A:Reference number: S04885
 A:Accession: S04885
 A:Molecule type: genomic RNA
 A:Residues: 1750-2164 <BAL>
 A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
 FEBS Lett. 247, 425-428, 1989
 A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
 A:Reference number: S03965; MUID:89232168; PMID:2541023
 A:Accession: S03965
 A:Molecule type: genomic RNA
 A:Residues: 1960-2164 <BAL2>
 A:Cross-references: EMBL:X15461
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein 1A #status predicted <C1A>
 F:1-27/Product: coat protein 1A #status predicted <C1A>
 F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-495/Product: coat protein 1C #status predicted <C1C>
 F:496-795/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: core protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 87.1%; Score 88; DB 1; Length 2230;
 Best Local Similarity 85.0%; Pred. No. 4.5e-06;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
 |||||
 Db 838 FSQAKISLFYTEHEIMKFS 857

RESULT 8
 F64300
 formate dehydrogenase (EC 1.2.1.2) alpha chain - Methanococcus jannaschii
 C:Species: Methanococcus jannaschii
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: F64300
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Kleink, H.P.; Fraser, C.M.; Smith, H.O.; Woese
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: F64300
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-378 <BUL>
 A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97987.1; PID:g1498763
 C:Genetics:
 A:Map position: REV8474-7338
 C:Keywords: oxidoreductase

Query Match 46.5%; Score 47; DB 2; Length 378;
 Best Local Similarity 47.1%; Pred. No. 4.7;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 AKISLFYTEHEIMKFS 20
 :||:| ||| ||:| :
 Db 174 SKVTIFNTEKEILKLN 190

RESULT 9
 H71869
 hypothetical protein jhp0940 - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
 C:Accession: H71869
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.
 ; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: H71869
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <ARN>
 A:Cross-references: GB:AE001523; GB:AE001439; NID:g4155513; PIDN:AAD06516.1; PID:g415
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: jhp0940
 C:Superfamily: Helicobacter pylori hypothetical protein jhp0940

Query Match 42.6%; Score 43; DB 2; Length 325;
 Best Local Similarity 60.0%; Pred. No. 19;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 5 KISLFYTEHEIMKF 19
 ||||| | :||
 Db 24 KISLFYNNELYMVF 38

RESULT 10

T22413
hypothetical protein F49C12.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22413
R:Gardner, A.
submitted to the EMBL Data Library, December 1995
A:Reference number: Z19562
A:Accession: T22413
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-410 <GLA>
A:Cross-references: EMBL:Z68227; PIDN:CAA92512.1; GSPDB:GN000022; CESP:F49C12.8
A:Experimental source: clone F49C12
C:Genetics:
A:Gene: CESP:F49C12.8
A:Map position: 4
A:Introns: 21/3; 302/3

Query Match 42.6%; Score 43; DB 2; Length 410;
Best Local Similarity 36.8%; Pred. No. 24;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKF 19
| : : : : : |
Db 156 FAMIRVGLFDFHLLINKF 174

RESULT 11
AB1747
glycine cleavage system protein H homolog lin2519 [imported] - *Listeria innocua* (strain
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AB1747
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1747
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97746.1; PID:gl6415041; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
Gene: lin2519
C:Superfamily: glycine cleavage system protein H; lipoyl/biotin-binding homology

Query Match 41.6%; Score 42; DB 2; Length 125;
Best Local Similarity 72.7%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LFYTEHEIMK 18
| : : : : : |
Db 7 LLYTEHEWVK 17

RESULT 12
A11377
glycine cleavage system protein H homolog lmo2425 [imported] - *Listeria monocytogenes* (s
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: A11377
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAD00503.1; PID:gl6411913; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo2425
C:Superfamily: glycine cleavage system protein H; lipoyl/biotin-binding homology

Query Match 41.6%; Score 42; DB 2; Length 125;
Best Local Similarity 72.7%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 LFYTEHEIMK 18
| : : : : : |
Db 7 LLYTEHEWVK 17

RESULT 13
TI8844
hypothetical protein C01H6.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: TI8844
R:Berks, M.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19030
A:Accession: TI8844
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-187 <WIL>
A:Cross-references: EMBL:Z71258; PIDN:CAA95778.1; GSPDB:GN00019; CESP:C01H6.3
A:Experimental source: clone C01H6
C:Genetics:
A:Gene: CESP:C01H6.3
A:Map position: 1
A:Introns: 42/3; 82/1; 109/1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein C01H6.3

Query Match 41.6%; Score 42; DB 2; Length 187;
Best Local Similarity 36.8%; Pred. No. 15;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKF 19
| : : : : : |
Db 161 FCNARISLQHAOSHOCVRF 179

RESULT 14
SI0842
gene ND5 intron 2 protein - *Neurospora crassa* mitochondrion
C:Species: mitochondrion *Neurospora crassa*
C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 21-Jul-2000
C:Accession: SI0842; SI0081
R:Nelson, M.A.; Macino, G.
Mol. Gen. Genet. 206, 307-317, 1987
A:Title: Structure and expression of the overlapping ND4L and ND5 genes of *Neurospora*
A:Reference number: S07320; MUID:87228330; PMID:3035337
A:Accession: SI0842
A:Molecule type: DNA
A:Residues: 1-545 <NEL>
A:Cross-references: EMBL:X05115
R:Almasan, A.; Mishra, N.C.
Genetics 120, 935-945, 1988
A:Title: Molecular characterization of the mitochondrial DNA of a new stopper mutant
A:Reference number: SI0081; MUID:89137935; PMID:2976009
A:Accession: SI0081
A:Molecule type: DNA
A:Residues: 457-545 <ALM>
A:Cross-references: EMBL:X14681; NID:gl4021; PIDN:CAA32812.1; PID:gl4022

A:Experimental source: strain Lindegren-Rockefeller RL3-8A

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC3

A:Introns: 108/3

C:Keywords: mitochondrion

Query Match 41.6%; Score 42; DB 2; Length 545;
Best Local Similarity 41.2%; Pred. No. 49;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 17

||:|:|:|:|:|:|:

Db 241 FSRAFLKLYMQEHPVL 257

RESULT 15

F88886

Protein F52G2.2 [imported] - Caenorhabditis elegans

A:Species: Caenorhabditis elegans

A:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: F88886

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: F88886

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1265 <STO>

A:Cross-references: GB:chr_IV; PIDN:CAB05207.1; PID:g3924776; GSPDB:GN00022; CESP:F52G2.

C:Genetics:

A:Gene: F52G2.2

A:Map position: 4

Query Match 41.6%; Score 42; DB 2; Length 1265;
Best Local Similarity 45.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20

||:|:|:|:|:|:|:

Db 1034 FSDQFVSAYVFELHEILATS 1053

Search completed: February 3, 2003, 13:15:36

Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:05:12 ; Search time 5.16667 Seconds
(without alignments)
160.554 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	100.0	2227	1	POLG_HPAVH
2	101	100.0	2227	1	POLG_HPAVH
3	101	100.0	2227	1	POLG_HPAVH
4	97	96.0	852	1	POLG_HPAVC
5	93	92.1	2226	1	POLG_HPAV2
6	93	92.1	2226	1	POLG_HPAV4
7	93	92.1	2226	1	POLG_HPAV8
8	88	87.1	2230	1	POLG_HPAVS
9	47	46.5	378	1	FDHA_METJA
10	43	42.6	410	1	PSD6_CAEEL
11	43	42.6	722	1	YC06_KLEPN
12	41.5	41.1	288	1	NIF3_YEAST
13	41	40.6	138	1	GCSH_PYPAB
14	41	40.6	304	1	META_THEMA
15	41	40.6	557	1	G6PI_ACICA
16	41	40.6	1040	1	MAN1_RAT
17	41	40.6	2431	1	POLN_SFV
18	40	39.6	418	1	B2AR_BOVIN
19	39	38.6	306	1	WK16_YEAST
20	39	38.6	380	1	TRPD_YEAST
21	39	38.6	534	1	UD15_HUMAN
22	39	38.6	634	1	YCX3_EUGER
23	39	38.6	707	1	HLX2_ECOLI
24	39	38.6	707	1	HLXB_ECOLI
25	39	38.6	707	1	HLXB_PROVU
26	39	38.6	837	1	LDLR_RABIT
27	39	38.6	860	1	LDLR_HUMAN
28	39	38.6	892	1	LDL2_XENLA
29	39	38.6	909	1	LDL1_XENLA
30	38	37.6	48	1	Z234_HAEIN
31	38	37.6	138	1	GCSH_PYPHO
32	38	37.6	265	1	CXBA_BACTW
33	38	37.6	317	1	RFAC_SALTY

34 38 37.6 319 1 RFAC_ECOLI
35 38 37.6 377 1 MTC3_CHVNI
36 38 37.6 396 1 REPA_BAGSU
37 38 37.6 400 1 ACH1_BOMMO
38 38 37.6 407 1 ARRH_LOEMI
39 38 37.6 448 1 RADA_HELPY
40 38 37.6 460 1 DB80_DROME
41 38 37.6 492 1 PDIL_SCHPO
42 38 37.6 533 1 MCP4_ECOLI
43 38 37.6 707 1 RT1B_ACPPL
44 38 37.6 720 1 WZC_ECO57
45 38 37.6 720 1 WZC_ECOLI

ALIGNMENTS

RESULT 1
POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P08617: P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
[2]
RC SEQUENCE FROM N.A.
RC STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
[3]
RC SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=8516289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -|- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

KW	Polyprotein; Coat protein; Core protein; Transferase;
KW	RNA-directed RNA polymerase;
FT	CHAIN 1 23
FT	CHAIN 24 245
FT	CHAIN 246 491
FT	CHAIN 492 794
FT	CHAIN 795 900
FT	CHAIN 901 1087
FT	CHAIN 1088 1422
FT	CHAIN 1423 1495
FT	CHAIN 1496 1518
FT	CHAIN 1519 1737
FT	CHAIN 1738 2226
SQ	SEQUENCE 2226 AA; 251152 MW; 6CD8A9106B4E2BF CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. NO. 2.2e-07;
Matches 18; Conservative 1; Mismatches 1; Indels

1 FSQAKISLFYTEEHEIMKFS 20
 834 FSOANISLFYTEEHEIMKFS 853

RESULT 6	
POLG_HPAAV4	
ID	POLG_HPAAV4 STANDARD; PRT; 2226 AA.
AC	P26581;
DT	01-AUG-1992 (Rel. 23, Created)
DT	01-AUG-1992 (Rel. 23, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE	P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE	P2D (EC 2.7.7.48)].
OS	Hepatitis A virus (strain 43c).
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC	Hepatovirus.
OX	NCBI_TaxID=12095;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=91162758; PubMed=1705995;
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA	Crombeans T., Jansen R.W.;
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus
RT	variants arising during persistent infection: evidence for genetic
RL	recombination.";
RL	J. Virol. 65:2056-2065(1991).

-I- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + [RNA](N).

-I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.

-I- PTM: SPECIFIC ENZYMAATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

-I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

EMBL; M59809; AAA5469.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR011205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
ET CHAIN 24 245 COAT PROTEIN VP2 (PIB).
LE

FT	CHAIN	246	491	COAT PROTEIN VP3 (P1C).
FT	CHAIN	492	794	COAT PROTEIN VP1 (P1D).
FT	CHAIN	795	900	CORE PROTEIN P2A.
FT	CHAIN	901	1087	CORE PROTEIN P2B.
FT	CHAIN	1088	1422	CORE PROTEIN P2C.
FT	CHAIN	1423	1495	PROBABLE PROTEIN P3A.
FT	CHAIN	1496	1518	PROBABLE PROTEIN P3B.
FT	CHAIN	1519	1737	PROBABLE PROTEIN P3C.
FT	CHAIN	1738	2226	RNA-DIRECTED POLYMERASE P3D.
SQ	SEQUENCE	2226	AA: 2511107	MM: 403B4CA80B09B75 CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. NO. 2.2e-07;
Matches 18; Conservative 1; Mismatches 1; Indels

Qy	1	FSQAKISLFYTEEHEIMKFS	20
Db	834	FSQANISLFYTEEHEIMKFS	853

```

RESULT 7
POLG_HPAV8
ID      POLG_HPAV8      STANDARD;      PRT;      2226 AA.
AC      P26582;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE      P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE      P3D (EC 2.7.7.48)].
DE      OS      Hepatitis A virus (strain 18f).
OS      OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC      OC      Hepatovirus.
OX      NCBI_TaxID=12096;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91162758; PubMed=1705995;
RA      Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA      Cromeans T., Jansen R.W.;
RT      "Antigenic and genetic variation in cytopathic hepatitis A virus
RT      variants arising during persistent infection: evidence for genetic
RT      recombination.";
RL      J. Virol. 65:2056-2065(1991).
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC      [RNA](N).
CC      -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC      EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC      VP3, AND VP4.
CC      -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

```

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59808; AAA45467.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PolyProtein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT

```

```
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. No. 2.2e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 834 FSQANISLFYTEHEIMKFS 853

RESULT 8
POLG_HPAYS STANDARD; PRT; 2230 AA.
ID POLG_HPAYS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232169; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapardze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00924; BAA00766.1; -
DR EMBL; X15461; CAA33490.1; -
DR PIR; A30470; GNNSA.
DR PIR; S04885; S04885.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PolyProtein; Coat protein; Core protein; Transferase.
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
```

```
FT CHAIN 1 27 COAT PROTEIN VP4 (P1A).
FT CHAIN 28 249 COAT PROTEIN VP2 (P1B).
FT CHAIN 250 495 COAT PROTEIN VP3 (P1C).
FT CHAIN 496 795 COAT PROTEIN VP1 (P1D).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 87.1%; Score 88; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 838 FSQAKISLFYTEHEIMKFS 857

RESULT 9
FDHA_METJA STANDARD; PRT; 378 AA.
ID FDHA_METJA STANDARD; PRT; 378 AA.
AC Q60314;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase alpha chain (EC 1.2.1.2).
GN MJ0006.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kleravage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN), ZINC, FAD, MAY BIND A 4FE-4S
CC CLUSTER (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67459; AAB97987.1; -
DR HSP; P07658; 1AA6.
DR TIGR; MJ0006; -
DR InterPro; IPR001467; Prok_Mboxred.
DR Pfam; PF00384; molybdopterin; 2.
DR PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
DR PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; 1.
DR PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
```

KW Hypothetical protein; Oxidoreductase; Zinc; Flavoprotein; Molybdenum;
FT FAD; NAD; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 8 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 11 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 15 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 378 AA; 42050 MW; E0C13060FF2609AA CRC64;

Query Match 46.5%; Score 47; DB 1; Length 378;
Best Local Similarity 47.1%; Pred. No. 1.8;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 AKISLFYTEHEIMKFS 20
: : : : I I I I I I :
Db 174 SKVTIFNTEKEILKLN 190

RESULT 10
PSD6_CAEEL STANDARD; PRT; 410 AA.
ID Q20585;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable 26S proteasome non-ATPase regulatory subunit 6.
GN F49C12.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
is involved in the ATP-dependent degradation of ubiquitinated
proteins (By similarity).
CC -!- DOMAIN: CONTAINS 1 PCI DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z68227; CAA92512.1; -;
CC WormPep; F49C12.8; GE03368.
CC InterPro; IPR000717; PCI.
CC Pfam; PF01399; PCI; 1.
CC SMART; SM00088; PIN1; 1.
CC Proteasome. 205 379 PCI.
FT DOMAIN
SQ SEQUENCE 410 AA; 47583 MW; F37390A3250109EE CRC64;

Query Match 42.6%; Score 43; DB 1; Length 410;
Best Local Similarity 36.8%; Pred. No. 9.2;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMKF 19
: : : : I I I I I I
Db 156 FAMIRVGLFFLDHLLINKF 174

RESULT 11
YC06_KLEPN
ID YC06_KLEPN STANDARD; PRT; 722 AA.
AC Q48452;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Putative tyrosine-protein kinase in cps region (EC 2.7.1.112) (ORF6).
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chedid;
RA MEDLINE=95204345; PubMed=7896702;
RA Arakawa Y., Wacharotayankun R., Nagatsuka T., Ito H., Kato N.,
Ohta M.;
RT "Genomic organization of the Klebsiella pneumoniae cps region
responsible for serotype K2 capsular polysaccharide synthesis in the
virulent strain Chedid.";
RL J. Bacteriol. 177:1788-1796(1995).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -!- PATHWAY: Exopolysaccharide biosynthesis.
CC -!- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC -!- SIMILARITY: BELONGS TO THE ETK/WZC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D21242; BAA04777.1; -;
CC InterPro; IPR003856; Wzz.
CC Pfam; PF02706; wzz; 1.
CC KW Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW Exopolysaccharide synthesis; Transmembrane; Inner membrane;
KW ATP-binding. 31 53 POTENTIAL.
FT TRANSMEM 427 449 POTENTIAL.
SQ SEQUENCE 722 AA; 80400 MW; 3CAD6910AE81C3D7 CRC64;

Query Match 42.6%; Score 43; DB 1; Length 722;
Best Local Similarity 44.4%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEIMK 18
: : : : I I I I I I
Db 326 FRESEISQLYTKHEPTYK 343

RESULT 12
NIF3_YEAST
ID NIF3_YEAST STANDARD; PRT; 288 AA.
AC P53081;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NGG1-interacting factor 3.
GN NIF3 OR YGL221C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
chromosome VII.";
RL Yeast 13:1077-1090(1997).
CC -!- SIMILARITY: BELONGS TO THE UPF0135 (NIF3) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -

Query Match	40.6%	Score 41;	DB 1;	Length 138;
Best Local Similarity	53.8%	Pred. No. 6;		

DE Glucose-6-phosphate isomerase (EC 5.3.1.9) (GPI) (Phosphoglucose

```

DE isomerase) (PGI) (Phosphohexose isomerase) (PHI).
GN PGI.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD4;
RA Stark M., Kaplan N., Ron E.; EMBL/GenBank/DBJ databases.
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate -> D-fructose 6-
CC phosphate.
CC -!- PATHWAY: Involved in glycolysis and in gluconeogenesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE GPI FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89900; CAA61993.1; -.
CC HSSP; Q9N1E2; 1HOX.
CC InterPro: IPR001672; G6P_Isomerase.
CC Pfam: PF00342; PGI; 1.
CC PRINTS: PR00662; G6PISOMERASE.
CC PROSITE; PS00765; P_GLUCOSE_ISOMERASE_1; 1.
CC PROSITE; PS00174; P_GLUCOSE_ISOMERASE_2; 1.
CC Isomerase; Gluconeogenesis; Glycolysis.
FT ACT_SITE 392 392 BY SIMILARITY.
FT ACT_SITE 520 520 BY SIMILARITY.
SQ SEQUENCE 557 AA; 63280 MW; D74AF214B139E4DC CRC64;

Query Match 40.6%; Score 41; DB 1; Length 557;
Best Local Similarity 50.0%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

Qy 1 FSQAKISLFYTEEHMKFS 20
Db 87 FSQNKIN--YTEQREAHWA 104

```

Search completed: February 3, 2003, 13:12:10
 Job time : 7.16667 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	101	100.0	112	12	Q86534	hepatitis a	
2	101	100.0	132	12	Q8V4n7	hepatitis a	
3	101	100.0	132	12	Q8V4N6	hepatitis a	
4	101	100.0	132	12	Q8V4N5	hepatitis a	
5	101	100.0	132	12	Q8V4N4	hepatitis a	
6	101	100.0	132	12	Q8V4N3	hepatitis a	
7	101	100.0	132	12	Q8V4N2	hepatitis a	
8	101	100.0	132	12	Q8V4N1	hepatitis a	
9	101	100.0	132	12	Q8V4M9	hepatitis a	
10	101	100.0	132	12	Q8V4M8	hepatitis a	
11	101	100.0	132	12	Q8V4M7	hepatitis a	
12	101	100.0	132	12	Q8V4M6	hepatitis a	
13	101	100.0	132	12	Q8V4M5	hepatitis a	
14	101	100.0	132	12	Q8V4M4	hepatitis a	
15	101	100.0	132	12	Q8V4M3	hepatitis a	
16	101	100.0	132	12	Q8V4M2	hepatitis a	

DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-21/10/94;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386846; AAL68507.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15288 MW; D8E1BF26DE38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
ID Q8V4N6 PRELIMINARY; PRT; 132 AA.
AC Q8V4N6
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-21/10/94;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386847; AAL68508.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
ID Q8V4N5 PRELIMINARY; PRT; 132 AA.
AC Q8V4N5
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-30/06/95;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386848; AAL68509.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15288 MW; D8E1BF26DE38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
ID Q8V4N4 PRELIMINARY; PRT; 132 AA.
AC Q8V4N4
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-28/07/95;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386849; AAL68510.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15288 MW; D8E1BF26DE38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
ID Q8V4N3 PRELIMINARY; PRT; 132 AA.
AC Q8V4N3
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.

```
RC STRAIN=LL-18/12/95;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386850; AAL68511.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 92 FSOAKISLFYTEEHEIMKFS 111

RESULT 7
Q8V4N2 PRELIMINARY; PRT; 132 AA.
ID Q8V4N2
AC Q8V4N2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-12/02/96;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386851; AAL68512.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15268 MW; D8FC83D50C31B87D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 92 FSOAKISLFYTEEHEIMKFS 111

RESULT 8
Q8V4N1 PRELIMINARY; PRT; 132 AA.
ID Q8V4N1
AC Q8V4N1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-17/06/97;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
```

```
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386852; AAL68513.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15354 MW; D05770901936C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 92 FSOAKISLFYTEEHEIMKFS 111

RESULT 9
Q8V4M9 PRELIMINARY; PRT; 132 AA.
ID Q8V4M9
AC Q8V4M9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-19/11/97;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386854; AAL68515.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEHEIMKFS 20
DB 92 FSOAKISLFYTEEHEIMKFS 111

RESULT 10
Q8V4M8 PRELIMINARY; PRT; 132 AA.
ID Q8V4M8
AC Q8V4M8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-05/12/97;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386855; AAL68516.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15302 MW; D8FC83D50C38287D CRC64;
```



```
RESULT 15
Q8V4M3
ID Q8V4M3 PRELIMINARY; PRT; 132 AA.
AC ORV4M3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-26/03/99;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jordi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
PL J. Gen. Virol. 82:2955-2963(2001).
PR EMBL; AF386860; AAL68521.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15288 MW; D8E1BF26DE38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FSQAKISLFYTEHEIMKFS 20
Db 92 FSQAKISLFYTEHEIMKFS 111
```

Search completed: February 3, 2003, 13:14:23
Job time : 22.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 : Search time 27.8333 seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNFPHGMLDEIAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	20	AAW42927	Immunogenic Hepati
2	104	100.0	21	AAW69444	Synthetic HAV P2A
3	104	100.0	2227	7 AAP60066	Sequence of viral
4	104	100.0	2227	11 AAR05697	Attenuated hepatit
5	104	100.0	2227	18 AAW34074	Hepatitis A virus
6	104	100.0	2227	21 AAB18607	Amino acid sequenc
7	104	100.0	2227	21 AAB18608	Amino acid sequenc
8	104	100.0	2227	21 AAB18609	Amino acid sequenc
9	104	100.0	2227	23 AAE19899	Hepatitis A virus
10	53	51.0	20	AAW42928	Immunogenic Hepati

11	53	51.0	21	22	AAW69445	Synthetic HAV P2A
12	47	45.2	387	23	ABB53661	Lactococcus lactis
13	45	43.3	195	22	AAU32447	Novel human secret
14	45	43.3	206	21	AAG32341	Arabisopsis thalia
15	45	43.3	242	21	AAG10097	Arabisopsis thalia
16	45	43.3	242	21	AAG32340	Arabisopsis thalia
17	45	43.3	383	21	AAG10096	Arabisopsis thalia
18	45	43.3	383	21	AAG32339	Arabisopsis thalia
19	45	43.3	399	21	AAG10095	Arabisopsis thalia
20	44	42.3	102	22	ABG11781	Novel human diagno
21	44	42.3	432	20	AA43167	Human succinyl-Co
22	44	42.3	432	21	AA43167	Human SC5H-2 prote
23	44	42.3	432	22	AAU32446	Novel human secret
24	43.5	41.8	267	21	AAG37721	Arabisopsis thalia
25	43.5	41.8	267	21	AAG37736	Arabisopsis thalia
26	43.5	41.8	269	21	AAG24210	Arabisopsis thalia
27	43.5	41.8	426	21	AAG37720	Arabisopsis thalia
28	43.5	41.8	426	21	AAG37735	Arabisopsis thalia
29	43.5	41.8	426	23	ABB93930	Herbicidally activ
30	43.5	41.8	428	21	AAG24209	Arabisopsis thalia
31	43.5	41.8	438	21	AAG37734	Arabisopsis thalia
32	43.5	41.8	440	21	AAG24208	Arabisopsis thalia
33	43.5	41.8	442	21	AAG37719	Arabisopsis thalia
34	43	41.3	98	19	AAW29645	Human secreted pro
35	43	41.3	98	20	AA439330	Protein product of
36	43	41.3	98	21	AAW44333	Human PRO725 prote
37	43	41.3	98	22	AAW49893	Human PRO725 prote
38	43	41.3	98	23	ABB95469	Human anglogenesis
39	43	41.3	98	23	ABB84863	Human PRO725 prote
40	43	41.3	98	23	AAU83644	Human PRO protein,
41	43	41.3	99	20	AA402719	Human secreted sec
42	43	41.3	110	20	AA436052	Extended human sec
43	43	41.3	116	22	AAW25843	Human protein sequ
44	43	41.3	124	18	AAW20410	H. pylori cytoplas
45	43	41.3	124	18	AAW24654	H. pylori cytoplas

ALIGNMENTS

RESULT 1
AAW42927
ID AAW42927 standard; peptide; 20 AA.
XX
AC AAW42927;
XX
DT 28-APR-1998 (first entry)
XX
XX Immunogenic Hepatitis A virus peptide YK-1327.
DE
XX
XX Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX Claim 18; Page 112; 140pp; English.

PS Peptides AAW42922-30 are immunogenic peptides corresponding to

CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are

CC substantially similar to a portion of the amino acid sequence of the P2A

CC protein of HAV corresponding to amino acids 792-980. The present peptide

CC is derived from amino acids 922-941, and has a reactivity of 31.3% with

CC acute sera. Compositions containing the peptides can be used to induce an

CC immune response to HAV in a mammal. The peptides can also be used to

CC detect the presence of antibodies against HAV in mammalian serum. The

CC peptides can also be used to make an antibody against HAV by

CC administering the peptide to a mammal.

XX

SQ Sequence 20 AA;

Query Match 100.0%; Score 104; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.3e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEETAANSKD 20

1 KVNFPHGMLDLEETAANSKD 20

Db

RESULT 2

AAB69444

ID AAB69444 standard; Peptide; 21 AA.

AC AAB69444;

XX

XX 20-APR-2001 (first entry)

XX

DE Synthetic HAV P2A peptide, SEQ ID NO: 44.

XX

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;

KW antigen; major structural capsid polypeptide; HAV antibody detection.

KW

XX Hepatitis A virus.

OS Synthetic.

OS

PN WO200105824-A2.

XX

XX 25-JAN-2001.

XX

XX 14-JUL-2000; 2000WO-US19267.

XX

XX 15-JUL-1999; 99US-0144412.

PR

(USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

XX Fields HA, Khudyakov YE;

XX

XX WPI; 2001-112681/12.

DR

XX

XX Synthetic peptides used as antigen sources for enzyme immunoassays

PT detecting anti-hepatitis A virus and as vaccines -

XX

XX Claim 13; Page 96; 130pp; English.

XX

XX The present sequence is one of a number of synthetic peptides which are

CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides

CC comprise antigenic epitopes of the major structural capsid polypeptides

CC or non-structural polypeptides of HAV with one or more glutamine

CC molecules at the carboxy end of the peptide. The peptides are used to

CC detect the presence of antibodies against HAV in mammalian serum, to

CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting

CC IgM antibodies in mammalian serum and detecting convalescence in a

CC mammal. The peptides are used to detect or quantify HAV antibodies in

CC samples in clinical or research-based assays using immunoblotting,

CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy

CC end of the peptides enhances the IgM antibody reactivity.

XX

SQ Sequence 21 AA;

Query Match 100.0%; Score 104; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.4e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEETAANSKD 20

1 KVNFPHGMLDLEETAANSKD 20

Db

RESULT 3

AAP60066

ID AAP60066 standard; Protein; 2227 AA.

XX

XX AAP60066;

XX

XX 26-JUN-1991 (first entry)

XX

DE Sequence of viral L434 polypeptide encoded by the complete

DE nucleotide sequence of the HAV genome.

XX

XX Diagnosis; vaccine; passive immunotherapy.

XX

OS Hepatitis A virus.

XX

XX

XX Key Location/Qualifiers

FT Region 1..245

FT /label= P1.1A

FT Region 246..491

FT /label= 1B

FT Region 492..836

FT /label= 1C

FT Region 837..980

FT /label= P2.2A

FT Region 981..1076

FT /label= 2B

FT Region 1077..1422

FT /label= 2C

FT Region 1423..1484

FT /label= P3.3A

FT Region 1485..1507

FT /label= 3B

FT Region 1508..1678

FT /label= 3C

FT Region 1679..2227

FT /label= 3D

XX

XX EP199480-A.

PN

XX

XX 29-OCT-1986.

PD

XX

XX 03-APR-1986; 86EP-0302465.

XX

XX 03-APR-1985; 85US-0719329.

PR

XX

XX (CHIR-) CHIRON CORP.

PA

XX

XX Dina D, Potter SJ, Vannest GA, Caput D;

PI

XX

XX WPI; 1986-286213/44.

DR

XX

XX N-PSDB; AAN60080.

DR

XX

XX Hepatitis A virus nucleotide sequence and polypeptide - and use

PT in prodn. of vaccines and diagnostic probes

XX

XX Claim 5; Fig 1; 18pp; English.

PS

XX

XX AAN60080 and oligonucleotide fragments are useful in detection of

several nucleotide changes distributed throughout the genome, and is attenuated for chimpanzees, elicits serum neutralising antibodies, and is suitable for use as an HAV vaccine. It is noted that not all the changes are necessary for attenuation and use as a vaccine.

CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is suitable for use as an HAV vaccine. It is noted that not all the changes CC are necessary for attenuation and use as a vaccine.

XX
SO Sequence 2227 AA:

sq	sequence	2227	aa;
Query Match	100.0%;	Score 104;	DB 11;
Best Local Similarity	100.0%;	Pred. NO. 2.4e-08;	Length 2227;
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps 0;			

QY 1 KVNPHGMLDLEEIAANSKD 20
|||||
Db 922 KVNPHGMLDLEEIAANSKD 941

RESULT 5
AAW34074
ID AAW34074 standard: Protein: 2227 AA.

AAW34074
ID AAW34074 standard; Protein; 2227 AA.
XX
AC AAW34074;

XX	AAW34074;
XX	AC
XX	XX
DT	27-APR-1998 (first entry)

XX
XX
DT 27-APR-1998 (first entry)
XX
XX
DE Hepatitis A virus HM-175 protein sequence.

XX Hepatitis A virus HM-175 protein sequence.
DE
XX
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW

XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
 KW infection; vaccine.
 XX

KW infection; vaccine.
XX
OS Hepatitis A virus HM-175.
XX

OS	Hepatitis A virus HM-175.
XX	
Key	Location/Qualifiers
FH	
ET	1..23

	Key	Location/Qualifiers
FH		
FT	Protein	1..23
FT		/label= VP4
FT	Protein	24...245

FT	/label= VP4
FT	24...245
FT	Protein
FT	/label= VP2
FT	246...491
FT	Protein

FT	/label= VP2
Protein	246..491
FT	/label= VP3
FT	492..791
Protein	

FT	/label= VP3
FT	492..791
Protein	/label= VP1
FT	792..980
FT	
Protein	

FT	/label= vpl
Protein	792..980
FT	/label= 2A
FT	981..1087

FT	/label= 2A
Protein	981..1087
FT	/label= 2B
FT	1088..1422
Protein	

FT	/label= 2B
FT	1088..1422
FT	/label= 2C
FT	1423..1496
FT	Protein

FT	/label= 2C
FT	1423..1496
Protein	
FT	/label= 3A
FT	1497..1519
Protein	

FT	/label= 3A
Protein	1497..1519
FT	/label= 3B
FT	1520..1738
Protein	

FT	/label= 3B
Protein	1520..1738
FT	/label= 3C
FT	1739..2227
Protein	

Protein	label = 3C
FT	1739..2227
FT	label = 3D
XX	

FT		/label= 3D
XX		
PN	W09740166-A2.	
XX		
XX		

PN	WO9740166-A2.
XX	
XX	
PD	30-OCT-1997.
XX	
XX	

PD	30-OCT-1997.
XX	
XX	
PF	18-APR-1997; 97WO-US06506.
XX	
XX	

PF 18-APR-1997; 97WO-US06506.
XX
PR 19-APR-1996; 96US-0015642.
XX

PR 19-APR-1996; 96US-0015642.
XX
XX (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

PI Emerson SU, Purcell RH, Raychaudhuri G;
XX WPI: 1997-535850/49.
DR N-PSDB; AAT93023.
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
PT as vaccines against HAV infection
XX
XX Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (I) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (I); (2) a cell transfected with (I) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (I) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (I) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX
XX Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFPHGMLDLEEAANSKD 20
DB 922 KVNFPHGMLDLEEAANSKD 941
RESULT 6
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
XX
XX AAB18607;
XX
XX 15-JAN-2001 (first entry)
XX
XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
XX Hepatitis A virus.
OS
XX US6113912-A.
XX
XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-0475886.
XX
XX 18-SEP-1992; 92US-0947338.
XX
XX 17-SEP-1993; 93WO-US08610.
XX
XX 10-MAR-1995; 95US-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI: 2000-586464/55.
XX
XX N-PSDB; AAA75476.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX

PS Disclosure; Fig 6A-K; 72pp; English.
XX
XX The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
XX Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFPHGMLDLEEAANSKD 20
DB 922 KVNFPHGMLDLEEAANSKD 941
RESULT 7
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
XX
XX AAB18608;
XX
XX 15-JAN-2001 (first entry)
XX
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.
DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX
XX P-35 virus.
KW
XX Hepatitis A virus.
OS
XX US6113912-A.
XX
XX 05-SEP-2000.
XX
XX 07-JUN-1995; 95US-0475886.
XX
XX 18-SEP-1992; 92US-0947338.
XX
XX 17-SEP-1993; 93WO-US08610.
XX
XX 10-MAR-1995; 95US-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI: 2000-586464/55.
XX
XX N-PSDB; AAA75477.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
XX Disclosure; Columns 67-78; 72pp; English.
XX
XX The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
XX Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFPHGMLDLEEAANSKD 20

Db 922 KVNPHGMLDLEETAANSKD 941
|||||

RESULT 8

AAE18609
ID AAE18609 standard; Protein; 2227 AA.

XX AC AAE18609;

DT 15-JAN-2001 (first entry)
XX DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
KW HAV 4380.

XX OS Hepatitis A virus.
XX PN US6113912-A.

PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX DR WPI: 2000-586464/55.

DR N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX Disclosure; Columns 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPHGMLDLEETAANSKD 20

Db 922 KVNPHGMLDLEETAANSKD 941
|||||

RESULT 9

AAE19899
ID AAE19899 standard; Protein; 2227 AA.

XX AC AAE19899;

DT 18-JUN-2002 (first entry)
XX DE Hepatitis A virus (HAV) protein.

XX Hepatitis A virus; HAV; Infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.

XX Hepatitis A virus.
OS WO200213855-A2.

PN 21-FEB-2002.

XX PF 15-AUG-2001; 2001WO-IB01808.

XX PR 17-AUG-2000; 2000US-225767P.

PR 29-AUG-2000; 2000US-229175P.

PR 03-NOV-2000; 2000US-0705547.

XX PA (TRIP-) TRIPEP AB.

XX PI Sallberg M, Hultgren C;

XX DR WPI: 2002-241837/29.

DR N-PSDB; AAD31766.

XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus -
XX Claim 11; Page 82-87; 120pp; English.

XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to
CC a hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.

XX The present sequence is hepatitis A virus (HAV) protein.

XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPHGMLDLEETAANSKD 20

Db 922 KVNPHGMLDLEETAANSKD 941
|||||

RESULT 10

AAW42928
ID AAW42928 standard; peptide; 20 AA.

XX AC AAW42928;

XX DT 28-APR-1998 (first entry)

XX DE Immunogenic Hepatitis A virus peptide YK-1328.

XX Immunogenic peptide; immunogenic epitope; P2A protein;

XX immune response; antibody.

XX Synthetic.

XX Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US06891.

XX 19-APR-1996; 96US-0015644.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

XX WPI; 1997-535831/49.
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal
XX Claim 18; Page 112; 140pp; English.
XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 931-950, and has a reactivity of 12.5% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX Sequence 20 AA;

Query Match 51.0%; Score 53; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 DLEETAAANSKD 20
DB 1 DLEETAAANSKD 11

RESULT 11
AAB69445
ID AAB69445 standard; Peptide; 21 AA.
XX AAB69445;
XX 20-APR-2001 (first entry)
XX Synthetic HAV P2A peptide, SEQ ID NO: 45.
DE Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
XX Hepatitis A virus.
OS Synthetic.
XX WO200105824-A2.

XX 25-JAN-2001.
XX 14-JUL-2000; 2000WO-US19267.
XX 15-JUL-1999; 99US-0144412.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Fields HA, Khudyakov YE;
PI WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX Claim 13; Page 97; 130pp; English.
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the Igm antibody reactivity.

XX Sequence 21 AA;

Query Match 51.0%; Score 53; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 DLEETAAANSKD 20
DB 1 DLEETAAANSKD 11

RESULT 12
ABB53661
ID ABB53661 standard; Protein; 387 AA.
XX ABB53661;
XX 16-MAY-2002 (first entry)
XX Lactococcus lactis protein recA.
DE Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
KW Lactococcus lactis IL1403.
XX Lactococcus lactis IL1403.
XX FR2807446-A1.
XX 12-OCT-2001.
XX 11-APR-2000; 2000FR-0004630.
XX 11-APR-2000; 2000FR-0004630.
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus
PT lactis and related species -
XX Claim 6; SEQ ID No 363; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 387 AA;

Query Match 45.2%; Score 47; DB 23; Length 387;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 KVNFPHGLDLEETAAANSKD 20

Db 340 KVRTAAGLLDEAEVAETTED 359

RESULT 13
AAU32447
ID AAU32447 standard; Protein: 195 AA.

XX AC AAU32447;

XX DT 18-DEC-2001 (first entry)

XX DE Novel human secreted protein #2938.

XX KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX OS Homo sapiens.

XX PN WO200179449-A2.

XX PD 25-OCT-2001.

XX PF 16-APR-2001; 2001WO-US08656.

XX PR 18-APR-2000; 2000US-0552929.

XX PR 26-JAN-2001; 2001US-0770160.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX DR WPI; 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -

XX PS Claim 20; Page 615; 765pp; English.

XX CC The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.

SQ Sequence 195 AA;

Query Match 43.3%; Score 45; DB 22; Length 195;
Best Local Similarity 42.1%; Pred. No. 14;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VNFPHGMLDLEETIAANSKD 20

Db 65 VGSPQGGVDIEEVAASNP 83

RESULT 14
AAG32341
ID AAG32341 standard; Protein: 206 AA.

XX

AC AAG32341;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38996.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 14-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139492.
XX PR 18-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 43.3%; Score 45; DB 21; Length 206;
Best Local Similarity 44.4%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NPHGMLDLLEETAAANSKD 20
:::|::|::|::|
Db 32 DYPVGLLDIEKILPPGKD 49

RESULT 15
AAG10097

ID AAG10097 standard; Protein; 242 AA.

XX AAG10097;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 8283.

KW Protein identification; signal transduction pathway; metabolic pathway;

hybridisation assay; genetic mapping; gene expression control; promoter;
 termination sequence.

Arabidopsis thaliana.
 EP1033405-A2.

06-SEP-2000.
 25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.
 05-MAR-1999; 99US-0123180.
 09-MAR-1999; 99US-0123548.
 23-MAR-1999; 99US-0125788.
 25-MAR-1999; 99US-0136264.
 29-MAR-1999; 99US-0136788.
 01-APR-1999; 99US-0127462.
 06-APR-1999; 99US-0128234.
 08-APR-1999; 99US-0128714.
 16-APR-1999; 99US-0129845.
 19-APR-1999; 99US-0130077.
 21-APR-1999; 99US-0130449.
 23-APR-1999; 99US-0130510.
 23-APR-1999; 99US-0130891.
 28-APR-1999; 99US-0131449.
 30-APR-1999; 99US-0132048.
 30-APR-1999; 99US-0132407.
 04-MAY-1999; 99US-0132488.
 05-MAY-1999; 99US-0132485.
 06-MAY-1999; 99US-0132486.
 06-MAY-1999; 99US-0132487.
 07-MAY-1999; 99US-0132863.
 11-MAY-1999; 99US-0134256.
 14-MAY-1999; 99US-0134218.
 14-MAY-1999; 99US-0134219.
 14-MAY-1999; 99US-0134221.
 14-MAY-1999; 99US-0134370.
 18-MAY-1999; 99US-0134768.
 19-MAY-1999; 99US-0134941.
 20-MAY-1999; 99US-0135124.
 21-MAY-1999; 99US-0135353.
 24-MAY-1999; 99US-0135629.
 25-MAY-1999; 99US-0136021.
 27-MAY-1999; 99US-0136392.
 28-MAY-1999; 99US-0136782.
 01-JUN-1999; 99US-0137222.
 03-JUN-1999; 99US-0137528.
 04-JUN-1999; 99US-0137502.
 07-JUN-1999; 99US-0137724.
 08-JUN-1999; 99US-0138094.
 10-JUN-1999; 99US-0138540.
 10-JUN-1999; 99US-0138847.
 14-JUN-1999; 99US-0139119.
 16-JUN-1999; 99US-0139452.
 16-JUN-1999; 99US-0139453.
 17-JUN-1999; 99US-0139492.
 18-JUN-1999; 99US-0139454.
 18-JUN-1999; 99US-0139455.
 18-JUN-1999; 99US-0139456.
 18-JUN-1999; 99US-0139457.
 18-JUN-1999; 99US-0139458.
 18-JUN-1999; 99US-0139459.
 18-JUN-1999; 99US-0139460.
 18-JUN-1999; 99US-0139461.
 18-JUN-1999; 99US-0139462.
 18-JUN-1999; 99US-0139463.
 18-JUN-1999; 99US-0139750.
 18-JUN-1999; 99US-0139763.
 21-JUN-1999; 99US-0139817.
 22-JUN-1999; 99US-0139899.
 23-JUN-1999; 99US-0140353.
 23-JUN-1999; 99US-0140354.

24-JUN-1999; 99US-0140695.
 28-JUN-1999; 99US-0140823.
 29-JUN-1999; 99US-0140991.
 30-JUN-1999; 99US-0141287.
 01-JUL-1999; 99US-0141842.
 01-JUL-1999; 99US-0142154.
 02-JUL-1999; 99US-0142055.
 06-JUL-1999; 99US-0142390.
 08-JUL-1999; 99US-0142803.
 09-JUL-1999; 99US-0142920.
 12-JUL-1999; 99US-0142977.
 13-JUL-1999; 99US-0143542.
 14-JUL-1999; 99US-0143624.
 15-JUL-1999; 99US-0144005.
 16-JUL-1999; 99US-0144085.
 16-JUL-1999; 99US-0144086.
 19-JUL-1999; 99US-0144325.
 19-JUL-1999; 99US-0144331.
 19-JUL-1999; 99US-0144332.
 19-JUL-1999; 99US-0144333.
 19-JUL-1999; 99US-0144334.
 19-JUL-1999; 99US-0144335.
 20-JUL-1999; 99US-0144352.
 20-JUL-1999; 99US-0144632.
 20-JUL-1999; 99US-0144884.
 21-JUL-1999; 99US-0144814.
 21-JUL-1999; 99US-0145086.
 21-JUL-1999; 99US-0145088.
 22-JUL-1999; 99US-0145085.
 22-JUL-1999; 99US-0145087.
 22-JUL-1999; 99US-0145089.
 22-JUL-1999; 99US-0145192.
 23-JUL-1999; 99US-0145145.
 23-JUL-1999; 99US-0145218.
 23-JUL-1999; 99US-0145224.
 26-JUL-1999; 99US-0145276.
 27-JUL-1999; 99US-0145913.
 27-JUL-1999; 99US-0145918.
 27-JUL-1999; 99US-0145919.
 28-JUL-1999; 99US-0145951.
 02-AUG-1999; 99US-0146386.
 02-AUG-1999; 99US-0146388.
 02-AUG-1999; 99US-0146389.
 03-AUG-1999; 99US-0147038.
 04-AUG-1999; 99US-0147204.
 04-AUG-1999; 99US-0147302.
 05-AUG-1999; 99US-0147192.
 05-AUG-1999; 99US-0147260.
 06-AUG-1999; 99US-0147303.
 06-AUG-1999; 99US-0147416.
 09-AUG-1999; 99US-0147493.
 09-AUG-1999; 99US-0147933.
 10-AUG-1999; 99US-0148171.
 11-AUG-1999; 99US-0148319.
 12-AUG-1999; 99US-0148341.
 13-AUG-1999; 99US-0148565.
 13-AUG-1999; 99US-0148684.
 16-AUG-1999; 99US-0149368.
 17-AUG-1999; 99US-0149175.
 18-AUG-1999; 99US-0149426.
 18-AUG-1999; 99US-0149722.
 20-AUG-1999; 99US-0149723.
 20-AUG-1999; 99US-0149929.
 23-AUG-1999; 99US-0149902.
 23-AUG-1999; 99US-0149930.
 25-AUG-1999; 99US-0150566.
 26-AUG-1999; 99US-0150884.
 27-AUG-1999; 99US-0151065.
 27-AUG-1999; 99US-0151066.
 27-AUG-1999; 99US-0151080.
 30-AUG-1999; 99US-0151303.
 31-AUG-1999; 99US-0151438.
 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 43.3%; Score 45; DB 21; Length 242;
Best Local Similarity 44.4%; Pred. No. 18;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
3 NFPHGMLDLEEIAANSKD 20
:::|::|::|
DB 68 DYPVGLLDIEKILPGKD 85

Search completed: February 3, 2003, 13:11:33
Job time : 29.8333 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 ; Search time 9.5 Seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNPHGMLDLEETAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	3	US-08-475-886-2
2	104	100.0	2227	3	US-08-475-886-4
3	104	100.0	2227	3	US-08-475-886-6
4	104	100.0	2227	4	US-08-397-232-2
5	104	100.0	2227	4	US-08-397-232-4
6	104	100.0	2227	4	US-09-171-387-2
7	104	100.0	2227	4	US-09-653-499-2
8	104	100.0	2227	4	US-09-653-499-4
9	104	100.0	2227	4	US-09-653-499-6
10	45	43.3	417	2	US-09-099-677A-6
11	45	43.3	417	3	US-09-261-471-6
12	44	42.3	432	2	US-09-099-677A-3
13	44	42.3	432	3	US-09-261-471-3
14	43	41.3	98	4	US-09-181-487-2
15	43	41.3	99	4	US-09-227-357-219
16	40	38.5	589	4	US-08-939-309-8
17	39	37.5	579	1	US-08-126-564A-31
18	39	37.5	579	5	PCT-US94-09143-31
19	39	37.5	609	4	US-09-115-475-16
20	39	37.5	609	4	US-09-115-475-19
21	39	37.5	609	4	US-09-115-475-22
22	39	37.5	609	4	US-09-115-475-25
23	39	37.5	609	4	US-09-115-475-28
24	39	37.5	609	4	US-09-115-475-31
25	39	37.5	1066	4	US-09-004-838-24
26	39	37.5	1323	4	US-09-004-838-90
27	39	37.5	1604	4	US-09-004-838-95

28	39	37.5	1805	4	US-09-004-838-92	Sequence 92, Appl
29	39	37.5	1817	4	US-09-004-838-125	Sequence 125, App
30	38	36.5	181	4	US-09-134-001C-3897	Sequence 3897, Ap
31	38	36.5	1250	3	US-08-938-291A-9	Sequence 9, Appli
32	38	36.5	1367	2	US-08-249-687C-2	Sequence 2, Appli
33	38	36.5	1367	2	US-08-625-819-2	Sequence 2, Appli
34	38	36.5	1367	3	US-08-746-559A-2	Sequence 2, Appli
35	38	36.5	1367	4	US-08-864-641B-18	Sequence 18, Appli
36	37.5	36.1	488	4	US-08-939-309-10	Sequence 10, Appli
37	37.5	36.1	568	4	US-09-238-373-2	Sequence 2, Appli
38	37.5	36.1	568	4	US-08-939-309-2	Sequence 2, Appli
39	37.5	36.1	568	4	US-08-939-309-4	Sequence 4, Appli
40	37	35.6	823	1	US-08-261-465-1	Sequence 1, Appli
41	37	35.6	823	1	US-08-405-254-5	Sequence 5, Appli
42	37	35.6	1525	4	US-09-396-651B-1	Sequence 1, Appli
43	36.5	35.1	350	2	US-08-828-922-1	Sequence 1, Appli
44	36.5	35.1	4544	1	US-08-469-486-52	Sequence 52, Appl
45	36.5	35.1	4544	2	US-08-469-658-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPHGMLDLEETAANSKD 20
Db 922 KVNPHGMLDLEETAANSKD 941

RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912

; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 4
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227

; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; -MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US97/06506
 ; FILING DATE: 18-APR-1997
 ; APPLICATION NUMBER: US60/015,642
 ; FILING DATE: 19-APR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: William S. Feller
 ; REGISTRATION NUMBER: 26,728
 ; REFERENCE/DOCKET NUMBER: 2026-4229US1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 758-4800
 ; TELEFAX: (212) 751-6849
 ; INFORMATION FOR SEQ ID NO: 2
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2227 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
 ; US-09-171-387-2

Query Match 100.0%; Score 104; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KVNFPHGMLDLEETAANSKD 20
 Db 922 KVNFPHGMLDLEETAANSKD 941

RESULT 7
 US-09-653-499-2
 ; Sequence 2, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
 ; US-09-653-499-2

Query Match 100.0%; Score 104; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KVNFPHGMLDLEETAANSKD 20
 Db 922 KVNFPHGMLDLEETAANSKD 941

RESULT 8
 US-09-653-499-4
 ; Sequence 4, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
 ; US-09-653-499-4

Query Match 100.0%; Score 104; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KVNFPHGMLDLEETAANSKD 20
 Db 922 KVNFPHGMLDLEETAANSKD 941

RESULT 9
 US-09-653-499-6
 ; Sequence 6, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: Attenuated (4380) HAV, strain HM-175
 ; US-09-653-499-6

Query Match 100.0%; Score 104; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KVNFPHGMLDLEETAANSKD 20
 Db 922 KVNFPHGMLDLEETAANSKD 941

RESULT 10
 US-09-099-677A-6
 ; Sequence 6, Application US/09099677A
 ; Patent No. 5965369
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Corley, Neil C.

Query Match 42.3%; Score 44; DB 2; Length 432;
Best Local Similarity 42.1%; Pred. No. 14;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VNFPHGMLDLEIEIAANSKD 20
Db 168 VGQPGGVDIIEVAASNPE 186

RESULT 13
US-09-261-471-3
; Sequence 3, Application US/09261471
; Patent No. 6025123
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/261,471
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/099,677
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0545 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSBPT06
; CLONE: 3273853
US-09-261-471-3

Query Match 42.3%; Score 44; DB 3; Length 432;
Best Local Similarity 42.1%; Pred. No. 14;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 2 VNFPHGMLDLEIEIAANSKD 20
Db 168 VGQPGGVDIIEVAASNPE 186

RESULT 14
US-09-181-487-2
; Sequence 2, Application US/09181487
; Patent No. 6165752
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID MALCOLM
; TITLE OF INVENTION: NOVEL COMPOUNDS

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestlia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/181,487
FILING DATE: 28-OCT-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9807082.4
FILING DATE: 01-APR-1998
APPLICATION NUMBER: 9815489.1
FILING DATE: 16-JUL-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestlia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30094
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-181-487-2

Query Match 41.3%; Score 43; DB 4; Length 98;
Best Local Similarity 44.4%; Pred. No. 3.7;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NFPHGMLDLEIEIAANSKD 20
Db 40 NLPEGVADLTQIDVNVQD 57

RESULT 15
US-09-227-357-219
; Sequence 219, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931

Job time : 10.5 secs

; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 219
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (99)
; OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-219

Query Match 41.3%; Score 43; DB 4; Length 99;
Best Local Similarity 44.4%; Pred. No. 3.7;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 3 NFPHGMLDLLEIAANSKD 20
| | | | | | | |
Db 40 NLPFGVADLTQIDVNVQD 57

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 ; Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNPHGMLDLEETAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	1 GNNVHM	genome polyprotein
2	104	100.0	2227	1 GNNYHR	genome polyprotein
3	104	100.0	2227	1 GNNYMK	genome polyprotein
4	104	100.0	2227	1 GNNYHB	genome polyprotein
5	101	97.1	2230	1 GNNYSA	genome polyprotein
6	47	45.2	333	2 A12485	hypothetical prote
7	47	45.2	387	2 B86669	RecA protein limpo
8	47	45.2	506	1 S58322	glycine-cRNA ligas
9	47	45.2	1060	2 F88710	protein ColG5.4 [1
10	47	45.2	1079	2 T30996	hypothetical prote
11	46	44.2	259	2 F82087	deoxyribose-phosph
12	46	44.2	444	2 C89768	conserved hypotet
13	46	44.2	677	2 H64574	DNA topoisomerase
14	45	43.3	195	2 T08812	probable succinate
15	45	43.3	383	2 F96362	hypothetical prote
16	45	43.3	417	2 A44529	succinate-CoA liga
17	45	43.3	852	2 B72685	hypothetical prote
18	45	43.3	935	2 T19011	hypothetical prote
19	45	43.3	1213	2 T19835	hypothetical prote
20	44	42.3	223	2 T37962	hypothetical prote
21	44	42.3	344	2 D97761	hypothetical prote
22	44	42.3	500	2 S50508	ANP1 protein - yea
23	43.5	41.8	373	2 A69773	hypothetical prote
24	43	41.3	140	2 E81659	conserved hypotet
25	43	41.3	141	2 G71501	hypothetical prote
26	43	41.3	257	2 A27592	2-deoxyribose-5-ph
27	43	41.3	259	2 H97374	2-deoxyribose-5-ph
28	43	41.3	309	2 H71089	hypothetical prote
29	43	41.3	327	2 T00876	probable glycerate

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: Genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: coat protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNPHGMLDLEETAANSKD 20

|||||

Db 922 KVNPHGMLDLEETAANSKD 941

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03903

R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 3
GNMYK
genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
R:Cohehen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: coat protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 4
GNMYB
genome polypeptide - human hepatitis A virus (strain MBH)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
R:Cohehen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: coat protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhard
C:Accession: J50303
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F:1-245/Product: coat protein 1A #status predicted <VP1>
F:246-491/Product: coat protein 1B #status predicted <VP2>
F:492-836/Product: coat protein 1C #status predicted <VP3>
F:837-980/Product: core protein 2A #status predicted <VP4>
F:981-1108/Product: core protein 2B #status predicted <VP5>
F:1109-1438/Product: core protein 2C #status predicted <VP6>
F:1439-1496/Product: protein 3A #status predicted <VP7>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VP8>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <VP9>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <VP10>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEIAANSKD 941

RESULT 5
GNMYA
genome polypeptide - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: JQ1080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; polypeptide
F:1-27/Product: coat protein 1A #status predicted <C1A>

Best Local Similarity 45.08; Pred. No. 6.7;

submitted to the EMBL Data Library, August 1999
A:Description: The sequence of *C. elegans* cosmid C01G5.
A:Reference number: Z20956
A:Accession: T30996
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1079 <BRA>
A:Cross-references: EMBL:U50068; PIDN:AAB37736.2
A:Experimental source: strain Bristol N2
C:Genetics:
A:Map position: IV
A:Introns: 10/1; 31/1; 75/1; 108/1; 144/1; 927/2; 1056/2
A:Note: C01G5.4

Query Match 45.2%; Score 47; DB 2; Length 1079;
Best Local Similarity 36.8%; Pred. No. 22;
Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPFGMLDLEIEIAANSK 20
: | | | | : | | | |
p 410 IGFSRGILDLKHVAGDARD 428

RESULT 11

F82087
deoxyribose-phosphate aldolase VC2350 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: F82087
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Paulsen, O.; Karp, P.D.; Klenk, H.P.; Aris, E.; White, O.; Kerlavage, A.R.; Sutton, G.G.; Fleischmann, R.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, N.; Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, A.; Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64574
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-677 <TOM>
A:Cross-references: GB:AE000559; GB:AE000511; NID:g2313536; PIDN:AAD07502.1; PID:g2313536
C:Superfamily: DNA topoisomerase I

Query Match 44.2%; Score 46; DB 2; Length 677;
Best Local Similarity 31.6%; Pred. No. 19;
Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEIEIAANSK 19
: | | | | : | | | | : | | | |
Db 346 RITHPHALKDLKVKCSDAK 364

RESULT 14

T08812
probable succinate-CoA ligase (GDP-forming) (EC 6.2.1.4) beta chain - human (fragment N)
N:Alternate names: protein DKFZp586M2023.1; succinyl-CoA synthetase (GDP-forming) bet
C:Species: *Homo sapiens* (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C:Accession: T08812
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
A:Reference number: Z16472
A:Reference number: Z16472
A:Accession: T08812
A:Molecule type: mRNA
A:Residues: 1-195 <ANS>
A:Cross-references: EMBL:AL050226
A:Experimental source: adult uterus; clone DKFZp586M2023
C:Genetics:
A:Note: DKFZp586M2023.1
C:Function:
A:Description: catalyzes the formation of succinyl-CoA from succinate with concomitant
C:Superfamily: succinate-CoA ligase (ADP-forming) beta chain
C:Keywords: acid-thiol ligase; coenzyme A; mitochondrion

Query Match 43.3%; Score 45; DB 2; Length 195;
Best Local Similarity 42.1%; Pred. No. 6.5;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPFGMLDLEIEIAANSK 20
: | | | | : | | | | : | | | |
Db 65 VGSPQGVDDIEEVAASNP 83

RESULT 15

conserved hypothetical protein SA0083 [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89768
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89768
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <KUR>
A:Cross-references: GB:BA000018; PID:g13700003; PIDN:BAB41302.1; GSPDB:GN00149
A:Experimental source: strain.N315

F96582
 hypothetical protein F1511.25 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F96582
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Saizberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F96582
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <STO>
 A:Cross-references: GB:AE005173; NID:g4587558; PIDN:AAD25789.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F1511.25
 A:Map position: 1

Query Match 43.3%; Score 45; DB 2; Length 383;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NPHGMLDLEETANSKD 20
 :| | | | | | | | | |
 Db 209 DYPVGLLDIEKILPPGKD 226

Search completed: February 3, 2003, 13:15:38
 Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:14:27 ; Search time 6 Seconds
(without alignments)
67.262 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNFPFGMLDLEETAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PublishedApplications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	9	US-10-104-966-12
2	104	100.0	2227	9	US-10-135-988-2
3	104	100.0	2227	9	US-10-135-988-4
4	104	100.0	2227	9	US-10-135-988-6
5	104	100.0	2227	10	US-09-929-955-12
6	43	41.3	98	9	US-09-978-295A-616
7	43	41.3	98	9	US-09-978-697-616
8	43	41.3	98	9	US-09-978-192A-616
9	43	41.3	98	9	US-09-999-832A-616
10	43	41.3	98	9	US-09-978-189-616
11	43	41.3	440	10	US-09-815-242-11399
12	43	41.3	609	10	US-09-815-242-14065
13	42	40.4	359	10	US-09-886-055-161
14	41	39.4	263	10	US-09-925-300-1112
15	40	38.5	596	10	US-09-797-039-8
16	39	37.5	121	10	US-09-867-550-134
17	39	37.5	321	10	US-09-815-242-14011
18	39	37.5	609	10	US-09-815-242-10399
19	39	37.5	993	10	US-09-815-242-5809

Sequence 12899, A
Sequence 13158, A
Sequence 154, App
Sequence 1856, Ap
Sequence 43252, A
Sequence 3, Appli
Sequence 1124, Ap
Sequence 10131, A
Sequence 1539, Ap
Sequence 11758, A
Sequence 120, App
Sequence 4, Appli
Sequence 2, Appli
Sequence 6422, Ap
Sequence 308, App
Sequence 285, App
Sequence 10650, A
Sequence 110, App
Sequence 1, Appli
Sequence 140, App
Sequence 142, App
Sequence 1362, Ap
Sequence 2, Appli
Sequence 44, Appl
Sequence 511, App
Sequence 581, App

39 37.5 1002 10 US-09-815-242-12899
39 37.5 1002 10 US-09-815-242-13158
38.5 37.0 514 9 US-09-712-363-154
38 36.5 60 10 US-09-867-550-1856
38 36.5 66 10 US-09-864-761-43252
38 36.5 375 10 US-09-996-606-3
38 36.5 451 10 US-09-764-864-1124
38 36.5 477 10 US-09-815-242-10131
38 36.5 561 10 US-09-764-864-1539
38 36.5 609 10 US-09-815-242-11758
38 36.5 1367 9 US-09-870-759-120
38 36.5 2835 10 US-09-885-535-4
37.5 36.1 568 9 US-09-740-369-2
37 35.6 256 9 US-09-738-626-6422
34 37 35.6 344 10 US-09-741-669-308
37 35.6 344 10 US-09-912-020-285
37 35.6 516 10 US-09-815-242-10650
37 35.6 542 12 US-10-078-929-110
37 35.6 1525 10 US-09-782-714-1
36.5 35.1 261 12 US-10-062-254-140
40 36.5 35.1 261 12 US-10-062-254-142
41 36.5 35.1 282 10 US-09-925-301-1362
42 36.5 35.1 4545 10 US-09-873-403-2
43 36 34.6 163 10 US-09-861-451A-44
44 36 34.6 186 10 US-09-841-132-511
45 36 34.6 193 10 US-09-841-132-581

ALIGNMENTS

RESULT 1

US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104.966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 104; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMLDLEETAANSKD 20

Db 922 KVNFPFGMLDLEETAANSKD 941

RESULT 2

US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20364262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 104; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
|||||
Db 922 KVNFPHGMLDLEEIAANSKD 941

RESULT 3
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20364262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 104; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
|||||
Db 922 KVNFPHGMLDLEEIAANSKD 941

RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20364262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 104; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
|||||
Db 922 KVNFPHGMLDLEEIAANSKD 941

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 104; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 5.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
|||||
Db 922 KVNFPHGMLDLEEIAANSKD 941

RESULT 6
US-09-978-295A-616
; Sequence 616, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742

; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;

Best Local Similarity 44.4%; Pred. No. 2.2;

Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPPHGMLEETAAANSKD 20
b 40 NLPEGVADLTQIDVNVQD 57

RESULT 7

US-09-978-697-616
; Sequence 616, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327

; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;
Best Local Similarity 44.4%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 3 NPHGMLDLIEIAANSKD 20
| | | | | | | | | |
Db 40 NLPEGVADLTQIDVNVQD 57

RESULT 8

US-09-978-192A-616
; Sequence 616, Application US/09978192A
; Patent No. US20020177553A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/065364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598

;
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;

Best Local Similarity 44.4%; Pred. No. 2.2;

Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NPHGMLDLEIAANSKD 20

||||:|:|:|:|:

Db 40 NLPEGVADLTQIDNVQD 57

RESULT 9

US-09-999-832A-616

; Sequence 616, Application US/09999832A

; Publication No. US20020192706A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C63

;
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195

; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07

; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 9; Length 98;
Best Local Similarity 44.4%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NFPHGMLEIEIAANSKD 20
Db 40 NLPEGVADLTQIDVNVQD 57
| | | | | | | | | |

RESULT 10

US-09-978-189-616
; Sequence 616, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249

1	PRIOR APPLICATION NUMBER: 60/081817
2	PRIOR FILING DATE: 1998-04-15
3	PRIOR APPLICATION NUMBER: 60/081819
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081952
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/081838
8	PRIOR FILING DATE: 1998-04-15
9	PRIOR APPLICATION NUMBER: 60/082568
10	PRIOR FILING DATE: 1998-04-21
11	PRIOR APPLICATION NUMBER: 60/082569
12	PRIOR FILING DATE: 1998-04-21
13	PRIOR APPLICATION NUMBER: 60/082704
14	PRIOR FILING DATE: 1998-04-22
15	PRIOR APPLICATION NUMBER: 60/082797
16	PRIOR FILING DATE: 1998-04-22
17	PRIOR APPLICATION NUMBER: 60/082796
18	PRIOR FILING DATE: 1998-04-23
19	PRIOR APPLICATION NUMBER: 60/083336
20	PRIOR FILING DATE: 1998-04-27
21	PRIOR APPLICATION NUMBER: 60/083322
22	PRIOR FILING DATE: 1998-04-28
23	PRIOR APPLICATION NUMBER: 60/083392
24	PRIOR FILING DATE: 1998-04-29
25	PRIOR APPLICATION NUMBER: 60/083495
26	PRIOR FILING DATE: 1998-04-29
27	PRIOR APPLICATION NUMBER: 60/083496
28	PRIOR FILING DATE: 1998-04-29
29	PRIOR APPLICATION NUMBER: 60/083499
30	PRIOR FILING DATE: 1998-04-29
31	PRIOR APPLICATION NUMBER: 60/083558
32	PRIOR FILING DATE: 1998-04-29
33	PRIOR APPLICATION NUMBER: 60/083558
34	PRIOR FILING DATE: 1998-04-29
35	PRIOR APPLICATION NUMBER: 60/083559
36	PRIOR FILING DATE: 1998-04-29
37	PRIOR APPLICATION NUMBER: 60/083500
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/083742
40	PRIOR FILING DATE: 1998-04-30
41	PRIOR APPLICATION NUMBER: 60/084366
42	PRIOR FILING DATE: 1998-05-05
43	PRIOR APPLICATION NUMBER: 60/084414
44	PRIOR FILING DATE: 1998-05-06
45	PRIOR APPLICATION NUMBER: 60/084441
46	PRIOR FILING DATE: 1998-05-06
47	PRIOR APPLICATION NUMBER: 60/084637
48	PRIOR FILING DATE: 1998-05-07
49	PRIOR APPLICATION NUMBER: 60/084600
50	PRIOR FILING DATE: 1998-05-07
51	PRIOR APPLICATION NUMBER: 60/085339
52	PRIOR FILING DATE: 1998-05-13
53	PRIOR APPLICATION NUMBER: 60/085338
54	PRIOR FILING DATE: 1998-05-13
55	PRIOR APPLICATION NUMBER: 60/085323
56	PRIOR FILING DATE: 1998-05-13
57	PRIOR APPLICATION NUMBER: 60/085582

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Search completed: February 3, 2003, 13:32:27
Job time : 7 secs

Qy 1 KVNPHGMLDLEEI 14
|:|:|: | | |:
Db 47 KINYPNTKLDPEQV 60

RESULT 14

US-09-925-300-1112
; Sequence 1112, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1112
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (19)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1112

Query Match 39.4%; Score 41; DB 10; Length 263;
Best Local Similarity 42.1%; Pred. No. 15;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KVNPHGMLDLEETAAASK 19
|:|:| | | | |:
Db 68 KVDFLKGMLQAEKLTSSSE 86

RESULT 15

US-09-797-039-8
; Sequence 8, Application US/09797039
; Patent No. US20020042099A1
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09/797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-797-039-8

Query Match 38.5%; Score 40; DB 10; Length 596;
Best Local Similarity 39.1%; Pred. No. 54;
Matches 9; Conservative 4; Mismatches 4; Indels 6; Gaps 1;

Qy 4 FPHGMLDL-----EEIAANSKD 20
| | | | | | |:
Db 267 FPHRMVELTGNVSSEFSMSKS 289

THIS PAGE BLANK (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:05:12 ; Search time 5.16667 Seconds
(without alignments)
160.554 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNFRHGMLDLEIAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	104	100.0	2226	1 POLG_HP4V4	P26581 hepatitis a
2	104	100.0	2226	1 POLG_HP4V8	P26582 hepatitis a
3	104	100.0	2227	1 POLG_HP4VH	P08617 hepatitis a
4	104	100.0	2227	1 POLG_HP4VL	P06441 hepatitis a
5	104	100.0	2227	1 POLG_HP4VM	P13901 hepatitis a
6	101	97.1	2230	1 POLG_HP4VS	P14553 simian hepa
7	100	96.2	2226	1 POLG_HP4V2	P26580 hepatitis a
8	47	45.2	387	1 REC2_LACLA	Q01840 lactococcus
9	47	45.2	404	1 SCB2_MOUSE	Q92218 mus musculus
10	47	45.2	505	1 SYG_THETH	P56206 thermus the
11	46	44.2	259	1 DEOC_VIBCH	Q9kpl7 vibrio chol
12	45	43.3	432	1 SCB2_HUMAN	Q96199 homo sapien
13	45	43.3	433	1 SCB2_PIG	P53590 sus scrofa
14	44	42.3	500	1 ANP1_YEAST	P32629 saccharomyc
15	43	41.3	98	1 SRG1_HUMAN	O75711 homo sapien
16	43	41.3	357	1 UBP2_CHICK	O57429 gallus gall
17	43	41.3	440	1 PUR8_HELPJ	Q92ka2 helicobacte
18	43	41.3	440	1 PUR8_HELPY	P56468 helicobacte
19	43	41.3	608	1 GLMS_SALTI	Q822q2 s glucosami
20	42	40.4	214	1 GLP1_PHANI	P45853 pharbitis n
21	42	40.4	290	1 YCQ6_YEAST	P25617 saccharomyc
22	41	39.4	193	1 YB71_HAEIN	P44339 haemophilus
23	41	39.4	218	1 Y363_RICPR	Q92d06 rickettsia
24	41	39.4	400	1 NIFS_ANAAZ	Q43884 anabaena az
25	41	39.4	400	1 NIFS_ANASP	P12623 anabaena sp
26	41	39.4	502	1 NU2C_MESVI	Q9muq6 mesostigma
27	41	39.4	552	1 SYQ_CLOPE	Q8xmp3 clostridium
28	40.5	38.9	237	1 VIB8_AGR75	P17798 agrobacteri
29	40.5	38.9	434	1 UD6_RICPR	O05973 rickettsia
30	40.5	38.9	1286	1 RPO1_VACCC	P20504 vaccinia vi
31	40.5	38.9	1286	1 RPO1_VARV	P33053 variola vir
32	40.5	38.9	1287	1 RPO1_VACCV	O07392 vaccinia vi
33	40	38.5	126	1 YF81_XYLFA	Q09p2c xylella fas

RESULT 1
POLG_HP4V4
ID POLG_HP4V4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination".
RL J. Virol. 65:2056-2065(1991).
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -|- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -|- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59809; AAA45469.1; --
DR MEROPS: C03.005; --
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.

P46905 bacillus su
O88623 mus musculu
O9jtx0 neisseria m
P47493 mycoplasma
P40486 saccharomyc
P37094 neisseria g
O75604 homo sapien
Q10364 schizosacch
P00882 escherichia
O8zv63 pyrobaculum
Q9rmn5 zymomonas m
O58175 methanococc

FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BF75 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEEIAANSKD 941

RESULT 2

ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 18f).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Cromeans T., Janssen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M59808; AAA45467.1; -
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).
FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).
FT CHAIN 492 794 COAT PROTEIN VP3 (P1C).
FT CHAIN 795 900 COAT PROTEIN VP1 (P1D).
FT CHAIN 901 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 PROBABLE PROTEIN P2C.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3A.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3B.
FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A6339C8D6B CRC64;

1

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEIAANSKD 20
|||||
DB 922 KVNPHGMLDLEEIAANSKD 941

RESULT 3

ID POLG_HPAVH STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain HM-175).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses."
RT J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M., Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus."
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr., Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase."
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate + [RNA](N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M14114; AAA45475.1; -
DR EMBL; M14707; AAA45465.1; -
DR EMBL; M14707; AAA45466.1; ALT_INIT.

EMBL; M16632; AAA45471.1; -
 PIR; A25981; GNNYHM.
 PIR; A25914; GNNYHM.
 PIR; A03905; A03905.
 MEROPS; C03.005; -
 InterPro; IPR000605; RNA_helicase.
 InterPro; IPR001205; RNA_pol_P3D.
 Pfam; PF00680; RNA_dep_RNA_pol; 1.
 Pfam; PF00910; RNA_helicase; 1.
 Polyprotein; Coat protein; Core protein; Transferase;
 RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1062 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 FT SEQUENCE 2227 AA; 251506 MW; 01E225E7AE740A6 CRC64;
 Query Match 100.0%; Score 104; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.4e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KVNPHGMLDLEIEAANSKD 20
 |||||
 DB 922 KVNPHGMLDLEIEAANSKD 941
 RESULT 4
 POLG_HPAPV STANDARD; PRT; 2227 AA.
 ID POLG_HPAPV P06441;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
 P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain LA).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
 Merryweather J., van Nest G., Dina D.;
 RT "Primary structure and gene organization of human hepatitis A virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 [RNA](N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M20273; AAA45474.1; -;
DR PIR; JS0303; GNNYHB.
DR MEROPS; C03.005; -;
DR InterPro; IPR000605; RNA_helicase.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 922 KVNPHGMLDLEEAANSKD 941
|||||

RESULT 6
POLG_HPAVS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapardze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.

CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D00924; BAA00766.1; -;
DR EMBL; X15461; CAA33490.1; -;
DR PIR; A30470; GNNYSA.
DR PIR; S04885; S04885.
DR MEROPS; C03.005; -;
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 97.1%; Score 101; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 4.7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDLEEAANSKD 20
DB 926 KVNPHGMLDLEEAANSKD 945
|||||

RESULT 7
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.


```

RL Nature 406:477-483(2000).
CC -|- CATALYTIC ACTIVITY: 2-deoxy-5-phosphate - D-
CC glyceraldhyde 3-phosphate + acetaldehyde.
CC -|- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES.
CC DEOC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE004305; AAF95493.1; -
CC TIGR; VC2350; -
CC InterPro; IPR002915; Deoc.
CC Pfam; PF01791; Deoc; 1.
CC Lyase; Schiff base; Complete proteome.
CC BINDING 166 166 SCHIFF-BASE (BY SIMILARITY).
CC SEQUENCE 259 AA; 27959 MW; 2876B07AC6527C73 CRC64;
CC -----
CC Query Match 44.2%; Score 46; DB 1; Length 259;
CC Best Local Similarity 52.9%; Pred. No. 1.6;
CC Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
CC
CC QY 3 NPPHGLMDLEEIAANSK 19
CC ||||| |.| |.|
CC Db 75 NPPHGNDDIEIAVETK 91
CC
CC
CC RESULT 12
CC SCB2_HUMAN STANDARD; PRT; 432 AA.
CC ID SCB2_HUMAN Q96199; O95195; Q8WUQ1;
CC AC Q96199; O95195; Q8WUQ1;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial precursor
CC DE (EC 6.2.1.4) [Succinyl-CoA synthetase, betaG chain] (Succinyl-CoA synthetase, beta subunit) (fragment).
CC DE SPECIFIC succinyl-CoA synthetase beta subunit (fragment).
CC DE SUGL2.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OC NCBI_TaxID=9606;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP TISSUE=Placenta, and Uterus;
CC RA Strausberg R.;
CC RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC [2]
CC RN SEQUENCE OF 29-432 FROM N.A.
CC RP TISSUE=Liver;
CC RC MEDLINE=98438536; PubMed=9765291;
CC RA Johnson J.D., Mehru J.G., Tews K., Milavetz B.I., Lambeth D.O.;
CC RA "Genetic evidence for the expression of ATP- and GTP-specific
CC RT succinyl-CoA synthetases in multicellular eucaryotes.";
CC RL J. Biol. Chem. 273:27580-27586(1998).
CC [3]
CC RN SEQUENCE OF 234-432 FROM N.A.
CC RP Mei G., Yu W., Gibbs R.A.;
CC RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +
CC phosphate.
CC -|- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC ACID CYCLE.
CC -|- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -|- SUBCELLULAR LOCATION: Mitochondrial.
CC -|- SIMILARITY: BELONGS TO THE SUCCINYL-COA SYNTHETASE BETA SUBUNIT
CC FAMILY.

```

CC CC PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC ACID CYCLE.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: BELONGS TO THE SUCCINYL-COA SYNTHETASE BETA SUBUNIT
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06944; AAA31120.1; ALT_INIT.
DR EMBL; L281187; CAB03559.1; -.
DR HSSP; P07460; ISCU.
DR InterPro; IPR003135; ATP-grasp.
DR Pfam; PF00549; CoA_ligase.
DR Pfam; PF02222; ATP-grasp; 1.
DR TIGRFAMs; TIGR01016; sucCoBeta; 1.
DR PROSITE; PS01217; SUCCINYL_COA_LIG_3; 1.
KW Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion;
KW Transit peptide.
FT NON_TER 1
FT TRANSIT <1 38 MITOCHONDRION.
FT CHAIN 39 433 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-
FT CHAIN 39 433 CHAIN.
FT FT
SQ SEQUENCE 433 AA; 46803 MW; AA04B72BC1B80E24 CRC64;

Query Match 43.3%; Score 45; DB 1; Length 433;
Best Local Similarity 42.1%; Pred. No. 4.3;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLEIEAANSKD 20
| | | | | : | | | | : | | : |
Db 169 VGSQGGVDIEEVAASNP 187

RESULT 14
ANPL_YEAST
ID ANPL_YEAST STANDARD; PRT; 500 AA.
AC P32629;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Aminotriphenyl propanediol resistance protein.
GN ANPL OR GEM3 OR YEL036C OR SYGP-ORF28.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-6441;
RX MEDLINE=94016558; PubMed=8411151;
RA Melnick L., Sherman F.;
RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
RT of Saccharomyces cerevisiae share a common ancestry.";
RL J. Mol. Biol. 233:372-388(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Sehl P., Komp C.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA "

RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mesedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95045382; PubMed=7957057;
RA Chapman R.E., Munro S.;
RT "The functioning of the yeast Golgi apparatus requires an ER protein
RT encoded by ANPL, a member of a new family of genes affecting the
RT secretory pathway.";
RL EMBL J. 13:4896-4907(1994).
CC -!- FUNCTION: INVOLVED IN THE ORGANIZATION OF THE SECRETORY PATHWAY.
CC REQUIRED TO MAINTAIN A FUNCTIONAL GOLGI APPARATUS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).
CC -!- SIMILARITY: BELONGS TO THE ANPL / MN9 / VAN1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65964; AAD13971.1; -.
DR EMBL; L22171; AAA34426.1; -.
DR EMBL; S66114; AAB28440.1; -.
DR EMBL; L22173; AAA34937.1; -.
DR EMBL; U18779; AAB65006.1; -.
DR PIR; S30846; S30846.
DR PIR; S38546; S38546.
DR SGD; S0000762; ANPL.
DR InterPro; IPR005109; Anpl.
DR Pfam; PF03452; Anpl; 1.
KW Glycoprotein; Transmembrane;
FT DOMAIN 1 15
FT TRANSMEM 16 27
FT DOMAIN 28 500
FT DOMAIN 446 467
FT DOMAIN 467
FT CONFLICT 220 224
FT CONFLICT 313 313
FT CONFLICT 472 500
FT TTRTKNLTKKPH (IN REF. 1).
FT PQGKPLDNDKNNKKHPKVPDLPDPRN -> RRGNLLMT
FT HHDKD -> QSQGQN (IN REF. 1).
FT F -> L (IN REF. 1).
SQ SEQUENCE 500 AA; 58182 MW; 845B395CE548CD14 CRC64;

Query Match 42.3%; Score 44; DB 1; Length 500;
Best Local Similarity 31.6%; Pred. No. 7.6;
Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 VNFPHGMLDLEIEAANSKD 20
: | | | | : | | | | : | | : |
Db 106 MTPPHNLIDLFLVSDSSD 124

RESULT 15
SRGL_HUMAN
ID SRGL_HUMAN STANDARD; PRT; 98 AA.
AC 075711;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Scapie-responsive protein 1 precursor (SCRG-1).
GN SCRG1
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98324999; PubMed=9660755;

RA Dron M., Dandoy-Dron F., Guillo F., Bendoudjema L., Hauw J.J.,
RA Lebon P., Dormont D., Tovey M.G.;
RT "Characterization of the human analogue of a Scrapie-responsive
RL gene";
RL J. Biol. Chem. 273:18015-18018(1998).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN THE CENTRAL NERVOUS
CC SYSTEM OF ADULT, BUT NOT AT ALL IN FETAL BRAIN. HIGH LEVELS OF
CC SCRG1 TRANSCRIPTS ARE ALSO OBSERVED IN TESTIS AND AORTA.
CC -1- SIMILARITY: BELONGS TO THE SCRG1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ224677; CAA12059.1; -;
DR MIM; 603163; -;
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 98 SCRAPIE-RESPONSIVE PROTEIN 1.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 98 AA; 11081 MW; A5F1FD40BF5401C0 CRC64;

Query Match 41.3%; Score 43; D8 1; Length 98;
Best Local Similarity 44.4%; Pred. No. 1.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NFPHGMLDLEEIAANSKD 20
| | | | | | | | | |
Db 40 NLPEGVADLTQIDVNVQD 57

Search completed: February 3, 2003, 13:12:12
Job time : 7.16667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:07:47 ; Search time 21.5 Seconds
(without alignments)
191.672 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNFPHGMLDLEETAAANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	104	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
2	104	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
3	104	100.0	251	12 Q9ENQ7	Q9enq7 hepatitis a
4	104	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
5	104	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
6	104	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
7	104	100.0	251	12 Q9ENQ1	Q9enq1 hepatitis a
8	104	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
9	104	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
10	104	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
11	104	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
12	104	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
13	104	100.0	251	12 Q9ENN7	Q9enn7 hepatitis a
14	104	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
15	104	100.0	251	12 Q9ENN5	Q9enn5 hepatitis a
16	104	100.0	251	12 Q9ENNA	Q9enn4 hepatitis a

17	104	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
18	104	100.0	1124	12 Q84780	Q84780 hepatitis a
19	104	100.0	1161	12 Q05794	Q05794 hepatitis a
20	104	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
21	104	100.0	2218	12 Q67824	Q67824 hepatitis a
22	104	100.0	2225	12 Q9DL32	Q9dl32 hepatitis a
23	104	100.0	2227	12 Q67825	Q67825 hepatitis a
24	104	100.0	2227	12 Q67826	Q67826 hepatitis a
25	104	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
26	104	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
27	104	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
28	104	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
29	104	100.0	2227	12 Q9IFH5	Q9ifh5 hepatitis a
30	104	100.0	2227	12 Q8VON6	Q8von6 hepatitis a
31	98	94.2	251	12 Q9ENR0	Q9enr0 hepatitis a
32	98	94.2	251	12 Q9ENQ8	Q9enq8 hepatitis a
33	98	94.2	251	12 Q9ENQ3	Q9enq3 hepatitis a
34	98	94.2	251	12 Q9ENQ0	Q9enq0 hepatitis a
35	98	94.2	251	12 Q9ENP8	Q9enp8 hepatitis a
36	98	94.2	251	12 Q9ENP6	Q9enp6 hepatitis a
37	98	94.2	251	12 Q9ENP3	Q9enp3 hepatitis a
38	98	94.2	251	12 Q9ENP0	Q9enp0 hepatitis a
39	98	94.2	251	12 Q9ENN8	Q9enn8 hepatitis a
40	98	94.2	251	12 Q9ENN3	Q9enn3 hepatitis a
41	98	94.2	2227	12 Q8QV03	Q8qv03 hepatitis a
42	97	93.3	184	12 Q87092	Q87092 simian hepa
43	95	91.3	2218	12 Q67817	Q67817 hepatitis a
44	94	90.4	251	12 Q9ENQ2	Q9enq2 hepatitis a
45	94	90.4	2227	12 Q9WM99	Q9wm99 hepatitis a

ALIGNMENTS

RESULT 1

Q9ENR1 ID Q9ENR1 PRELIMINARY; PRT; 251 AA.
AC Q9ENR1; 2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A1;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1;
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 58A520D873893445 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAAANSKD 20
|
Db 86 KVNFPHGMLDLEETAAANSKD 105

RESULT 2

Q9ENQ9 ID Q9ENQ9 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

```
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A159;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEEIAANSKD 20
|||||
86 KVNFPHGMLDLEEIAANSKD 105

Db

RESULT 3
Q9ENQ7
ID Q9ENQ7 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ7;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A161;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047656; BAB12164.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28766 MW; A0EC02E3609C4D5B CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEEIAANSKD 20
|||||
86 KVNFPHGMLDLEEIAANSKD 105

Db

RESULT 4
Q9ENQ6
ID Q9ENQ6 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ6;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A162;
RA Fujiwara K.;
```

```
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEEIAANSKD 20
|||||
86 KVNFPHGMLDLEEIAANSKD 105

Db

RESULT 5
Q9ENQ5
ID Q9ENQ5 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ5;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A20;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEEIAANSKD 20
|||||
86 KVNFPHGMLDLEEIAANSKD 105

Db

RESULT 6
Q9ENQ4
ID Q9ENQ4 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ4;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A201;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPHGMLDLEEIAANSKD 20
|||||
86 KVNFPHGMLDLEEIAANSKD 105

Db
```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAANSKD 20
 |||||

Db 86 KVNFPHGMLDLEETAANSKD 105

RESULT 7

Q9ENQ1 ID Q9ENQ1 PRELIMINARY; PRT; 251 AA.

AC Q9ENQ1; SEQUENCE FROM N.A.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A206;
 RA Fujiwara K.;
 RT "Hepatitis A virus."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB047662; BAB12170.1; -.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAANSKD 20
 |||||

Db 86 KVNFPHGMLDLEETAANSKD 105

RESULT 8

Q9ENP9 ID Q9ENP9 PRELIMINARY; PRT; 251 AA.

AC Q9ENP9; SEQUENCE FROM N.A.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A303;
 RA Fujiwara K.;
 RT "Hepatitis A virus."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB047664; BAB12172.1; -.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CA5C1A CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAANSKD 20
 |||||

Db 86 KVNFPHGMLDLEETAANSKD 105

RESULT 9

Q9ENP7

ID Q9ENP7 PRELIMINARY; PRT; 251 AA.
 AC Q9ENP7; SEQUENCE FROM N.A.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A306;
 RA Fujiwara K.;
 RT "Hepatitis A virus."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB047666; BAB12174.1; -.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAANSKD 20
 |||||

Db 86 KVNFPHGMLDLEETAANSKD 105

RESULT 10

Q9ENP5 ID Q9ENP5 PRELIMINARY; PRT; 251 AA.

AC Q9ENP5; SEQUENCE FROM N.A.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A407;
 RA Fujiwara K.;
 RT "Hepatitis A virus."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB047668; BAB12176.1; -.
 FT NON_TER 1
 FT NON_TER 251
 SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
 Best Local Similarity 100.0%; Pred. No. 8.5e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KVNFPHGMLDLEETAANSKD 20
 |||||

Db 86 KVNFPHGMLDLEETAANSKD 105

RESULT 11

Q9ENP2 ID Q9ENP2 PRELIMINARY; PRT; 251 AA.

AC Q9ENP2; SEQUENCE FROM N.A.
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.

```
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
Db 86 KVNFPHGMLDLEEIAANSKD 105

RESULT 12
Q9ENN7
ID Q9ENN7 PRELIMINARY; PRT; 251 AA.
AC Q9ENN7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
Db 86 KVNFPHGMLDLEEIAANSKD 105

RESULT 13
Q9ENN7
ID Q9ENN7 PRELIMINARY; PRT; 251 AA.
AC Q9ENN7;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A712;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047676; BAB12184.1; -.
FT NON_TER 1
FT NON_TER 251
```

```
FT NON_TER 251
SQ SEQUENCE 251 AA; 28785 MW; BFE79D3A26134F18 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
Db 86 KVNFPHGMLDLEEIAANSKD 105

RESULT 14
Q9ENN6
ID Q9ENN6 PRELIMINARY; PRT; 251 AA.
AC Q9ENN6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A713;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
Db 86 KVNFPHGMLDLEEIAANSKD 105

RESULT 15
Q9ENN5
ID Q9ENN5 PRELIMINARY; PRT; 251 AA.
AC Q9ENN5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A75;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047678; BAB12186.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28658 MW; 98E8EED00B2EDF10 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPHGMLDLEEIAANSKD 20
Db 86 KVNFPHGMLDLEEIAANSKD 105
```


Mon Feb 3 13:52:46 2003

Db 86 KVNPHGMLDLEETAANSKD 105

Search completed: February 3, 2003, 13:14:23
Job time : 21.5 secs

us-09-171-432a-44.rspt

Page 5

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 ; Search time 27.8333 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEETAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	20	AAW42928	Immunogenic Hepati
2	101	100.0	21	AAW69445	Synthetic HAV P2A
3	101	100.0	2277	AAP00066	Sequence of viral
4	101	100.0	2227	AAW05697	Attenuated hepatit
5	101	100.0	2227	AAW34074	Hepatitis A virus
6	101	100.0	2227	AAW18607	Amino acid sequenc
7	101	100.0	2227	AAW18608	Amino acid sequenc
8	101	100.0	2227	AAW18609	Amino acid sequenc
9	101	100.0	2227	AAW19899	Hepatitis A virus
10	53	52.5	20	AAW42927	Immunogenic Hepati

11	53	52.5	21	22	AAW69444	Synthetic HAV P2A
12	46	45.5	424	22	ABW71556	Drosophila melanog
13	44.5	44.1	1712	22	ABW60536	Drosophila melanog
14	44	43.6	80	23	ABW02822	Human ORFX protein
15	43	42.6	152	18	AAW55492	H. pylori ORF hp2e
16	43	42.6	157	18	AAW55327	H. pylori ORF hp2p
17	43	42.6	350	20	AAW27183	(S)-3'-hydroxy-N-m
18	43	42.6	536	21	AAW48090	Arabidopsis thalia
19	43	42.6	1201	20	AAW90345	Drosophila sp. Cos
20	43	42.6	1201	22	ABW58431	Drosophila melanog
21	42	41.6	89	22	AAW59799	Propionibacterium
22	42	41.6	177	21	AAW48507	Arabidopsis thalia
23	42	41.6	226	21	AAW48506	Arabidopsis thalia
24	42	41.6	265	21	AAW48505	Arabidopsis thalia
25	42	41.6	365	18	AAW13493	Pentaerythritol te
26	42	41.6	365	20	AAW29461	Enterobacter cloac
27	42	41.6	402	22	ABW60260	Drosophila melanog
28	42	41.6	425	22	ABW09060	Thermus caldophilu
29	42	41.6	445	22	AAW40763	Human polypeptide
30	42	41.6	445	22	AAW40764	Human polypeptide
31	42	41.6	445	22	AAW40765	Human polypeptide
32	42	41.6	538	21	AAW81713	Streptococcus pneu
33	42	41.6	641	22	AAW38978	Human polypeptide
34	42	41.6	679	23	AAW17073	Human r2 protein.
35	42	41.6	698	22	AAW38977	Human polypeptide
36	42	41.6	705	22	AAW39032	Human polypeptide
37	42	41.6	705	22	AAW94135	Human protein sequ
38	42	41.6	706	22	AAW38979	Human polypeptide
39	42	41.6	1837	21	AAW85564	Human homologue of
40	41	40.6	72	22	ABW43574	Peptide #11080 enc
41	41	40.6	72	22	ABW26533	Protein #8532 enco
42	41	40.6	72	22	AAW64511	Human brain expres
43	41	40.6	72	22	AAW77319	Human bone marrow
44	41	40.6	72	22	AAW21245	Peptide #7679 enco
45	41	40.6	72	22	AAW37471	Peptide #11508 enc

ALIGNMENTS

RESULT 1
AAW42928
ID AAW42928 standard; peptide; 20 AA.
XX
AC AAW42928;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1328.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
PI
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX PS Claim 18; Page 112; 140pp; English.

XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to

CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are

CC substantially similar to a portion of the amino acid sequence of the P2A

CC protein of HAV corresponding to amino acids 792-980. The present peptide

CC is derived from amino acids 931-950, and has a reactivity of 12.5% with

CC acute sera. Compositions containing the peptides can be used to induce an

CC immune response to HAV in a mammal. The peptides can also be used to

CC detect the presence of antibodies against HAV in mammalian serum. The

CC peptides can also be used to make an antibody against HAV by

CC administering the peptide to a mammal.

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 101; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 9.5e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 DLEETAAANSKDFPNMSETDL 20

1 DLEETAAANSKDFPNMSETDL 20

1 DLEETAAANSKDFPNMSETDL 20

RESULT 2

AAB69445

ID AAB69445 standard; Peptide; 21 AA.

AC AAB69445;

XX 20-APR-2001 (first entry)

XX Synthetic HAV P2A peptide, SEQ ID NO: 45.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;

KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO2000105824-A2.

XX 25-JAN-2001.

XX 14-JUL-2000; 2000WO-US19267.

XX 15-JUL-1999; 99US-0144412.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khuyakov YE;

XX WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays

XX detecting anti-hepatitis A virus and as vaccines

XX Claim 13; Page 97; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are

CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides

CC comprise antigenic epitopes of the major structural capsid polypeptides

CC or non-structural polypeptides of HAV with one or more glutamine

CC molecules at the carboxy end of the peptide. The peptides are used to

CC detect the presence of antibodies against HAV in mammalian serum, to

CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting

CC IGM antibodies in mammalian serum and detecting convalescence in a

CC mammal. The peptides are used to detect or quantify HAV antibodies in

CC samples in clinical or research-based assays using immunoblotting,

CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy

CC end of the peptides enhances the Igm antibody reactivity.

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 101; DB 22; Length 21;

Best Local Similarity 100.0%; Pred. No. 1e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20

1 DLEETAAANSKDFPNMSETDL 20

DB 1 DLEETAAANSKDFPNMSETDL 20

RESULT 3

AAP60066

ID AAP60066 standard; Protein; 2227 AA.

XX AAP60066;

XX 26-JUN-1991 (first entry)

XX Sequence of viral L434 polypeptide encoded by the complete

DE nucleotide sequence of the HAV genome.

XX Diagnosis; vaccine; passive immunotherapy.

XX Hepatitis A virus.

XX Key Location/Qualifiers

FT Region 1..245

FT /label= P1.1A

FT Region 246..491

FT /label= 1B

FT Region 492..836

FT /label= 1C

FT Region 837..980

FT /label= P2.2A

FT Region 981..1076

FT /label= 2B

FT Region 1077..1422

FT /label= 2C

FT Region 1423..1484

FT /label= P3.3A

FT Region 1485..1507

FT /label= 3B

FT Region 1508..1678

FT /label= 3C

FT Region 1679..2227

FT /label= 3D

XX EPI99480-A.

XX 29-OCT-1986.

XX 03-APR-1986; 86EP-0302465.

XX 03-APR-1985; 85US-0719329.

XX (CHIR-) CHIRON CORP.

XX Dina D, Potter SJ, Vannest GA, Caput D;

XX WPI; 1986-286213/44.

XX N-PSDB; AAN60080.

XX Hepatitis A virus nucleotide sequence and polypeptide - and use

XX in prodn. of vaccines and diagnostic probes

XX Claim 5; Fig 1; 18pp; English.

XX AAN60080 and oligonucleotide fragments are useful in detection of

CC hepatitis A virus; transformed hosts may be used for expression of
CC polypeptides and fragments useful in vaccines without risk of
CC infection by the virus or in prodn. of particles which are capable
CC of inducing immunocompetent B cells for passive immunotherapy. Pref.
CC epitope is derived from AAs 445-657 or 792-848 of the HAV
CC polypeptide sequence (AAP60066).

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 7; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
|||||

Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 4

AAW05697
D AAR05697 standard; protein; 2227 AA.

XX
AC AAR05697;

DT 15-AUG-1990 (first entry)

DE Attenuated hepatitis A virus.

KW Hepatitis A virus; vaccine; attenuated.

OS Hepatitis A virus, strain HM-175.

XX Key Location/Qualifiers

FT Region 1..23
/label=VP4 - 1A

FT Region 24..245
/label=VP2 - 1B

FT Region 246..491
/label=VP3 - 1C

FT Region 492..791
/label=VP1 - 1D

FT Region 792..980
/label=2A

FT Region 981..1087
/label=2B

FT Region 1088..1422
/label=2C

FT Region 1423..1496
/label=3A

FT Region 1497..1519
/label=3B - VPg

FT Region 1520..1738
/label=3C

FT Region 1739..2227
/label=3D

XX US4894228-A.

XX 16-JAN-1990.

PD 12-JUL-1988; 88US-0217824.

PR 12-JUL-1988; 88US-0217824.

PR 12-JUL-1988; 88US-0652967.

XX (USSH) US DEPT HEALTH & HUMAN.

XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstone SM;

PI Daemer RJ, Gust ID;

XX WPI; 1990-075557/10.

DR N-PSDB; AAQ03512.

XX

PT Vaccine against hepatitis A virus infection - comprises novel
PT attenuated hepatitis A virus strain.

PS Claim 1; Fig 1; 18pp; English.

XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
|||||

Db 931 DLEETAAANSKDFPNMSETDL 950

RESULT 5

AAW34074
ID AAW34074 standard; Protein; 2227 AA.

XX
AC AAW34074;

DT 27-APR-1998 (first entry)

DE Hepatitis A virus HM-175 protein sequence.

KW HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW infection; vaccine.

OS Hepatitis A virus HM-175.

XX Key Location/Qualifiers

FT Protein 1..23
/label= VP4

FT Protein 24..245
/label= VP2

FT Protein 246..491
/label= VP3

FT Protein 492..791
/label= VP1

FT Protein 792..980
/label= 2A

FT Protein 981..1087
/label= 2B

FT Protein 1088..1422
/label= 2C

FT Protein 1423..1496
/label= 3A

FT Protein 1497..1519
/label= 3B

FT Protein 1520..1738
/label= 3C

FT Protein 1739..2227
/label= 3D

XX WO9740166-A2.

PN 30-OCT-1997.

XX 18-APR-1997; 97WO-US06506.

PR 19-APR-1996; 96US-0015642.

XX (USSH) US SEC DEPT HEALTH.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Emerson SU, Purcell RH, Raychaudhuri G;
 XX WPI; 1997-535850/49.
 DR N-PSDB; AAT93023.
 XX
 PT Human attenuated HAV genome containing simian HAV 2C gene - useful
 PT as vaccines against HAV infection
 XX
 XX
 PS Disclosure; Fig 13A-D; 66pp; English.
 XX
 CC This protein sequence is encoded by the human hepatitis A virus
 CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
 CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
 CC kidney cells. A claimed DNA construct (I) comprises a genome of
 CC HAV, where the genome is a human attenuated HAV genome in which a
 CC region of the 2C gene has been replaced by a corresponding region
 CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
 CC region of the 2C gene from AGM-27 contained in the construct
 CC preferably encodes amino acids 120-328 of the 2C protein, amino
 CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
 CC transcript of (I); (2) a cell transfected with (I) or the RNA
 CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
 CC HAV of (3); and (5) a host cell containing the HAV of (3). (I) or
 CC its RNA transcript, can be used as a vaccine for preventing HAV in
 CC a mammal. (I) or the RNA transcript can also be used to stimulate
 CC the production of protective antibodies in the mammal.
 XX
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 18; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEEIAANSKDFPNMSETDL 20
 DB 931 DLEEIAANSKDFPNMSETDL 950
 RESULT 6
 AAB18607
 ID AAB18607 standard; Protein; 2227 AA.
 XX
 AC AAB18607;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
 XX
 XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX
 PD 05-SEP-2000.
 XX
 PF 07-JUN-1995; 95US-0475886.
 PR 18-SEP-1992; 92US-0947338.
 PR 17-SEP-1993; 93WO-US08610.
 PR 10-MAR-1995; 95US-0397232.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX
 DR WPI; 2000-586464/55.
 DR N-PSDB; AAA75476.
 XX
 PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type -
 XX

PS Disclosure; Fig 6A-K; 72pp; English.
 XX
 CC The present sequence is derived from a wild type hepatitis A virus
 CC (HAV) strain HM-174. The sequence is modified to produce HAV which
 CC are adapted to growth in the human fibroblast-like cell line MRC-5.
 CC The HAV is able to propagate in MRC-5 cells and retain appropriate
 CC attenuation. It is useful as a live vaccine for prophylaxis of
 CC hepatitis A in humans and other primates.
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 21; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEEIAANSKDFPNMSETDL 20
 DB 931 DLEEIAANSKDFPNMSETDL 950
 RESULT 7
 AAB18608
 ID AAB18608 standard; Protein; 2227 AA.
 XX
 AC AAB18608;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
 XX
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW P-35 virus.
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX
 PD 05-SEP-2000.
 XX
 PF 07-JUN-1995; 95US-0475886.
 PR 18-SEP-1992; 92US-0947338.
 PR 17-SEP-1993; 93WO-US08610.
 PR 10-MAR-1995; 95US-0397232.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX
 DR WPI; 2000-586464/55.
 DR N-PSDB; AAA75477.
 XX
 PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type -
 XX
 PS Disclosure; Columns 67-78; 72pp; English.
 XX
 CC The present sequence is derived from passage 35 of a wild type
 CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
 CC designated P-35 virus. The sequence is modified to produce HAV which
 CC are adapted to growth in the human fibroblast-like cell line MRC-5.
 CC The HAV is able to propagate in MRC-5 cells and retain appropriate
 CC attenuation. It is useful as a live vaccine for prophylaxis of
 CC hepatitis A in humans and other primates.
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 21; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DLEEIAANSKDFPNMSETDL 20

```
Db 931 DLEEIAANSKDFPNMSETDL 950
|||||
RESULT 8
AAB18609
ID AAB18609 standard; Protein; 2227 AA.
XX AAB18609;
XX AC
XX 15-JAN-2001 (first entry)
XX DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX HAV 4380.
XX OS Hepatitis A virus.
XX US6113912-A.
XX 05-SEP-2000.
XX 07-JUN-1995; 95US-0475886.
XX 18-SEP-1992; 92US-0947338.
XX 17-SEP-1993; 93WO-US08610.
XX 10-MAR-1995; 95US-0397232.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI: 2000-586464/55.
XX N-PSDB; AAB75478.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type -
XX Disclosure; Columns 93-104; 72pp; English.
XX The present sequence is derived from a live attenuated hepatitis A
XX virus (HAV) of the invention, designated HAV 4380. The sequence is
XX produced by modifying wild type HAV strain HM-174. The HAV of the
XX invention are adapted to growth in the human fibroblast-like cell
XX line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
XX appropriate attenuation. It is useful as a live vaccine for prophylaxis
XX of hepatitis A in humans and other primates.
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950
RESULT 10
AAW42927
ID AAW42927 standard; peptide; 20 AA.
XX AC
XX AAW42927;
XX 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1327.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein;
XX immune response; antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX WO9740147-A1.
XX 30-OCT-1997.
XX 18-APR-1997; 97WO-US06891.
XX 19-APR-1996; 96US-0015644.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
Hepatitis A virus.
WO200213855-A2.
21-FEB-2002.
15-AUG-2001; 2001WO-IB01808.
17-AUG-2000; 2000US-225767P.
29-AUG-2000; 2000US-229175P.
03-NOV-2000; 2000US-0705547.
(TRIP-) TRIPEP AB.
Sallberg M, Hultgren C;
WPI: 2002-241837/29.
N-PSDB; AAD31766.
Vaccine compositions for treating and preventing disease, preferably
hepatitis C virus infection, comprises ribavirin and antigen that has
epitope present in hepatitis C virus -
Claim 11; Page 82-87; 120pp; English.
The invention relates to a composition comprising ribavirin and an
antigen preferably non structural 3 protein (NS3)/4A fragment of
hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
sequence. The composition is useful for enhancing an immune response to
a hepatitis C antigen in humans, domestic, sport or pet species and as
a vaccines for treating and preventing HCV infections. The composition is
also useful for treating viral, bacterial, fungal diseases and cancer.
The present sequence is hepatitis A virus (HAV) protein.
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950
RESULT 10
AAW42927
ID AAW42927 standard; peptide; 20 AA.
XX AC
XX AAW42927;
XX 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1327.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein;
XX immune response; antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX WO9740147-A1.
XX 30-OCT-1997.
XX 18-APR-1997; 97WO-US06891.
XX 19-APR-1996; 96US-0015644.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
```

XX WPI; 1997-535831/49.
XX Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal
XX Claim 18; Page 112; 140pp; English.
XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 922-941, and has a reactivity of 31.3% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
XX Sequence 20 AA;
Query Match 52.5%; Score 53; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKD 11
DB 10 DLEETAAANSKD 20
RESULT 11
AAB69444
ID AAB69444 standard; Peptide; 21 AA.
XX AAB69444;
AC AAB69444;
DT 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 44.
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
XX 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
PR 15-JUL-1999; 99US-0144412.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 96; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of

CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
XX Sequence 21 AA;
Query Match 52.5%; Score 53; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETAAANSKD 11
DB 10 DLEETAAANSKD 20
RESULT 12
AAB71556
ID AAB71556 standard; Protein; 424 AA.
XX AAB71556;
AC AAB71556;
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 41460.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL15659.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 41460; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 424 AA;
Query Match 45.5%; Score 46; DB 22; Length 424;

Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNKSE 17
I: | | | | | | | | | |
Db 180 DVAELAASRDMPHMAK 196

RESULT 13

ABBS0536
ID ABB60536 standard; Protein; 1712 AA.

XX AC ABB60536;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 8400.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

XX W0200171042-A2.

PN 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL04639.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 8400; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1712 AA;

Query Match 44.1%; Score 44.5; DB 22; Length 1712;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 DLEE-IAANSKDFPNSETD 19
I: | | | | | | | | | |
Db 44 DIQDLIAETGGLPNSSETD 63

RESULT 14

ABP02822
ID ABP02822 standard; Protein; 80 AA.

XX AC ABP02822;

XX 24-JUN-2002 (first entry)
DE Human ORFX protein sequence SEQ ID NO:5626.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.

XX Homo sapiens.

XX W0200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

PR 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI; 2002-106308/14.

DR N-PSDB; ABN18574.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders -

XX Disclosure; SEQ ID 5626; 1037pp; English.

XX The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 80 AA;

Query Match 43.6%; Score 44; DB 23; Length 80;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 EEIAANSKDFPNMSET 18

Db 18 QDIASDAKDFTNPET 33

Db 16 VALNAKDFSKTSEDL 31

Search completed: February 3, 2003, 13:11:35
Job time : 29.8333 secs

RESULT 15
AAW55492
ID AAW55492 standard; Protein; 152 AA.
XX
XX
AC AAW55492;
XX
XX 30-JUN-1998 (first entry)
XX
XX H. pylori ORF hp6el0967_24889750_f2_7 secreted protein.
XX
XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW bacterium.
XX
XX
OS Helicobacter pylori.
XX
XX W09737044-A1.
XX
XX 09-OCT-1997.
XX
XX 27-MAR-1997; 97WO-US05223.
XX
XX 06-DEC-1996; 96US-0761318.
PR 29-MAR-1996; 96US-0625811.
PR 02-APR-1996; 96US-0758731.
PR 25-OCT-1996; 96US-0736905.
PR 28-OCT-1996; 96US-0738859.
XX
XX (ASTR) ASTRA AB.
XX
XX Alm RA, Smith D;
XX
XX WPI; 1997-503122/46.
DR N-PSDB; AAV24901.
XX
XX Helicobacter pylori nucleic acid sequences and encoded
PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
XX
XX Claims 14,94; Page 698; 1145pp; English.
XX
XX This sequence is a H. pylori secreted protein.
CC The protein may be used in a vaccine to prevent or treat H. pylori
CC infection or to identify H. pylori polypeptide binding compounds,
CC useful as potential H. pylori life cycle activators or inhibitors.
CC The DNA and probes derived from it may be used for the
CC identification of H. pylori in a sample, and the diagnosis of
CC H. pylori infection. Nucleic acid sequences complementary to the
CC DNA act as antisense sequences, and can be used to prevent the
CC translation of H. pylori mRNA. Antibodies against the protein can
CC be used in immunoassays to evaluate the abundance and distribution
CC of H. pylori-specific antigens. The genomic sequence of H. pylori
CC (ATCC 55679) was determined from overlapping contigs generated
CC by mechanically shearing the bacterial DNA. The sequences were
CC analysed for ORF of at least 180 nucleotides, and the predicted
CC coding regions defined by computer evaluation. To identify likely
CC H. pylori antigens for vaccine development, the amino acid
CC sequences predicted from various ORF were analysed for significant
CC homology to other known or exported membrane proteins. Having
CC identified and determined the sequences of interest, particular
CC regions can be isolated from H. pylori by PCR amplification for
CC recombinant polypeptide production, e.g. in E. coli hosts.
XX
XX Sequence 152 AA;

Query Match 42.6%; Score 43; DB 18; Length 152;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 IAANSKDFPNMSETDL 20

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 ; Search time 9.5 Seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEIAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	2227	3	US-08-475-886-2
2	101	100.0	2227	3	US-08-475-886-4
3	101	100.0	2227	3	US-08-475-886-6
4	101	100.0	2227	4	US-08-397-232-2
5	101	100.0	2227	4	US-08-397-232-4
6	101	100.0	2227	4	US-09-171-387-2
7	101	100.0	2227	4	US-09-653-499-2
8	101	100.0	2227	4	US-09-653-499-4
9	101	100.0	2227	4	US-09-653-499-6
10	44	43.6	907	3	US-08-938-830-26
11	44	43.6	907	3	US-09-020-222-26
12	43	42.6	1201	4	US-09-098-901-2
13	42	41.6	365	2	US-08-983-352-2
14	41	40.6	376	4	US-09-200-965-2
15	39.5	39.1	224	4	US-09-134-001C-3648
16	39	38.6	50	1	US-08-127-351-14
17	39	38.6	50	1	US-08-480-367B-14
18	39	38.6	50	1	US-08-487-221A-14
19	39	38.6	50	1	US-08-480-370-14
20	39	38.6	168	4	US-09-134-001C-4664
21	39	38.6	352	4	US-09-286-691-26
22	39	38.6	352	4	US-09-687-147-26
23	39	38.6	387	4	US-09-314-847A-4
24	39	38.6	387	4	US-09-570-778A-2
25	39	38.6	387	4	US-09-570-778A-7
26	39	38.6	387	4	US-09-570-778A-8
27	39	38.6	387	4	US-09-570-778A-9

28	39	38.6	396	1	US-08-430-024-2	Sequence 2, Appl
29	39	38.6	396	1	US-08-782-009-2	Sequence 2, Appl
30	39	38.6	396	3	US-09-017-302-2	Sequence 2, Appl
31	39	38.6	984	1	US-08-242-932-2	Sequence 2, Appl
32	39	38.6	984	1	US-08-714-481-2	Sequence 2, Appl
33	39	38.6	984	5	PCT-US95-06111-2	Sequence 2, Appl
34	39	38.6	1098	4	US-08-923-992A-8	Sequence 8, Appl
35	39	38.6	1104	4	US-08-923-992A-4	Sequence 4, Appl
36	39	38.6	1128	4	US-08-923-992A-6	Sequence 6, Appl
37	39	38.6	1164	4	US-08-923-992A-2	Sequence 2, Appl
38	39	38.6	1164	4	US-08-923-992A-10	Sequence 10, Appl
39	38	37.6	99	1	US-08-004-492-9	Sequence 9, Appl
40	38	37.6	99	2	US-08-484-905-60	Sequence 60, Appl
41	38	37.6	99	3	US-08-481-985B-60	Sequence 60, Appl
42	38	37.6	99	4	US-08-370-476-60	Sequence 73, Appl
43	38	37.6	105	1	US-08-081-539-73	Sequence 73, Appl
44	38	37.6	105	1	US-08-466-647-73	Sequence 73, Appl
45	38	37.6	114	4	US-09-347-819-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEIAANSKDFPNMSETDL 20
Db 931 DLEIAANSKDFPNMSETDL 950

RESULT 2

US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feller
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNMSETDL 20
|||||
Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 10
US-08-938-830-26
; Sequence 26, Application US/08938830
; Patent No. 6040437
; GENERAL INFORMATION:
; APPLICANT: Lasky, Laurence A.
; APPLICANT: Dowbenko, Donald J.
; TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage

; TITLE OF INVENTION: Furrow-Associated Proteins (PSTPIPS)
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,830
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/798419
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1066P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-938-830-26

Query Match 43.6%; Score 44; DB 3; Length 907;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSET 18
Db 49 LQELAASSADIPEVGST 65

RESULT 11
US-09-020-222-26
; Sequence 26, Application US/09020222
; Patent No. 6111073
; GENERAL INFORMATION:
; APPLICANT: Lasky, Laurence A.
; TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,222
; FILING DATE: 06-Feb-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/798419
; FILING DATE: 02/07/1997
; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1066r1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-020-222-26

Query Match 43.6%; Score 44; DB 3; Length 907;
Best Local Similarity 47.1%; Pred. No. 35;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSET 18
Db 49 LQELAASSADIPEVGST 65

RESULT 12
US-09-098-901-2
; Sequence 2, Application US/09098901B
; Patent No. 6218144
; GENERAL INFORMATION:
; APPLICANT: Scott, Matthew
; APPLICANT: Sisson, John C.
; TITLE OF INVENTION: Costal2 Genes and their Uses
; FILE REFERENCE: SUN-65P
; CURRENT APPLICATION NUMBER: US/09/098,901B
; CURRENT FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: 60/051,347
; EARLIER FILING DATE: 1997-06-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1201
; TYPE: PRT
; ORGANISM: D. Melanogaster
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1201)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-098-901-2

Query Match 42.6%; Score 43; DB 4; Length 1201;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 ETAANSKDFPNMSETD 19
Db 48 QVPAAEKDNPSISSTD 63

RESULT 13
US-08-983-352-2
; Sequence 2, Application US/08983352
; Patent No. 5928859
; GENERAL INFORMATION:
; APPLICANT: Nicklin, Stephen
; APPLICANT: Binks, Peter R.
; APPLICANT: Bruce, Neil C.
; APPLICANT: French, Christopher E.
; TITLE OF INVENTION: DETECTION AND BIODEGRADATION OF
; TITLE OF INVENTION: EXPLOSIVES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5928859th Glebe Rd.
; CITY: Arlington
; STATE: VA

```

1  COUNTRY: USA
2  ZIP: 22201-4741
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: Floppy disk
5  COMPUTER: IBM PC compatible
6  OPERATING SYSTEM: PC-DOS/MS-DOS
7  SOFTWARE: PatentIn Release #1.0, Version #1.30
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER: US/08/983.352
10 FILING DATE: 02-FEB-1998
11 CLASSIFICATION: 435
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: PCT/GB96/01629
14 FILING DATE: 08-JUL-1996
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: GB 9514138.8
17 FILING DATE: 11-JUL-1995
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: GB 9607220.2
20 FILING DATE: 04-APR-1996
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Crawford, Arthur R.
23 REGISTRATION NUMBER: 25,327
24 REFERENCE/DOCKET NUMBER: 124-617
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 703-816-4000
27 TELEFAX: 703-816-4100
28 INFORMATION FOR SEQ ID NO: 2:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 365 amino acids
31 TYPE: amino acid
32 TOPOLOGY: linear
33 MOLECULE TYPE: protein
34 US-08-983-352-2
35
36 Query Match 41.6%; Score 42; DB 2; Len
37 Best Local Similarity 47.4%; Pred. No. 26;
38 Matches 9; Conservative 4; Mismatches 6;
39
40 QY 2 LEIAAKSDFPNMSDTL 20
41 :||:| : :|||
42
43 Db 258 IEELAKRGIALHMSDTL 276
44
45 RESULT 14
46 US-09-200-965-2
47 : Sequence 2, Application US/09200965
48 : Patent No. 6133422
49 GENERAL INFORMATION:
50 APPLICANT: ROSEN, ET AL.
51 TITLE OF INVENTION: Thrombin Inhibitor
52 NUMBER OF SEQUENCES: 4
53 CORRESPONDENCE ADDRESS:
54 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
55 ADDRESSES: CECCHI, STEWART & OLSTEIN
56 STREET: 6 BECKER FARM ROAD
57 CITY: ROSELAND
58 STATE: NEW JERSEY
59 COUNTRY: USA
60 ZIP: 07068
61 COMPUTER READABLE FORM:
62 MEDIUM TYPE: 3.5 INCH DISKETTE
63 COMPUTER: IBM PS/2
64 OPERATING SYSTEM: MS-DOS
65 SOFTWARE: WORD PERFECT 5.1
66 CURRENT APPLICATION DATA:
67 APPLICATION NUMBER: US/09/200.965
68 FILING DATE:
69 CLASSIFICATION:
70 PRIOR APPLICATION DATA:
71 APPLICATION NUMBER: 08/171.817
72 FILING DATE: 22-DECEMBER-1993
73 ATTORNEY/AGENT INFORMATION:

```

```
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-47
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: LINEAR
; TOPOLOGY: PROTEIN
; MOLECULE TYPE: PROTEIN
US-09-200-965-2

Query Match          40.6%; Score 41; 1
Best Local Similarity 33.3%; Pred. No. 3
Matches 10; Conservative 3; Mismatch

QY      1 DLEIIAAN-----SKDFPNMSETDL 20
       1 : : | : | : |
Db      281 DMESVLRLNGMTAFELGKADFGSMSTDL 310
              ||| |||||

RESULT 15
US-09-134-001C-3648
; Sequence 3648, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: EPIDERMIS FOR DIAGN
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3648
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3648

Query Match          39.1%; Score 39.5; 1
Best Local Similarity 52.9%; Pred. No. 3
Matches 9; Conservative 3; Mismatch

QY      2 LEEIAANSKDF-PNMSE 17
       1 |||| | : | || :
Db     169 LEEILAEKGKEDFNLHQ 195

Search completed: February 3, 2003, 13:16:43
Job time : 11.5 secs
```

Search completed: February 3, 2003, 13:16:43
Job time : 11.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:14:27 ; Search time 6 Seconds
(without alignments)
67.262 Million cell updates/sec

Title: US-09-171-432A-45
Perfect score: 101
Sequence: 1 DLEETAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	2227	9 US-10-104-966-12	Sequence 12, Appli
2	101	100.0	2227	9 US-10-135-988-2	Sequence 2, Appli
3	101	100.0	2227	9 US-10-135-988-4	Sequence 4, Appli
4	101	100.0	2227	9 US-10-135-988-6	Sequence 6, Appli
5	101	100.0	2227	10 US-09-929-955-12	Sequence 12, Appl
6	41	40.6	72	10 US-09-864-761-41831	Sequence 14831, A
7	41	40.6	459	10 US-09-925-300-1440	Sequence 1440, Ap
8	40	39.6	284	10 US-09-810-997-1	Sequence 1, Appli
9	40	39.6	381	10 US-09-815-242-11280	Sequence 11280, A
10	40	39.6	499	10 US-09-864-761-35385	Sequence 35385, A
11	39	38.6	331	10 US-09-803-286A-10	Sequence 10, Appl
12	39	38.6	341	10 US-09-815-242-10657	Sequence 10657, A
13	39	38.6	387	9 US-10-037-677-4	Sequence 4, Appli
14	39	38.6	396	10 US-09-815-242-10387	Sequence 10387, A
15	39	38.6	417	10 US-09-815-242-10492	Sequence 10492, A
16	39	38.6	420	9 US-09-764-868-912	Sequence 912, App
17	39	38.6	978	10 US-09-893-817-4	Sequence 4, Appli
18	39	38.6	978	10 US-09-893-817-8	Sequence 8, Appli
19	39	38.6	1111	10 US-09-815-242-12955	Sequence 12955, A

20 39 38.6 1247 10 US-09-803-286A-2 Sequence 2, Appli
21 38 37.6 212 10 US-09-854-280-14 Sequence 14, Appl
22 38 37.6 212 10 US-09-854-208-14 Sequence 14, Appl
23 38 37.6 286 10 US-09-764-898-256 Sequence 256, App
24 38 37.6 375 10 US-09-996-606-3 Sequence 3, Appli
25 38 37.6 888 9 US-09-738-626-5302 Sequence 5302, Ap
26 37 37.6 1170 10 US-09-801-368-298 Sequence 298, App
27 37.5 37.1 1136 10 US-09-815-242-12447 Sequence 12447, A
28 37.5 37.1 1136 10 US-09-815-242-12814 Sequence 12814, A
29 37.5 37.1 1179 10 US-09-815-242-5522 Sequence 5522, Ap
30 37 36.6 104 9 US-09-738-626-5488 Sequence 5488, Ap
31 37 36.6 158 10 US-09-860-352A-8 Sequence 8, Appli
32 37 36.6 176 9 US-09-984-245-191 Sequence 191, App
33 37 36.6 200 9 US-09-991-496-84 Sequence 84, Appl
34 37 36.6 200 10 US-09-874-923-84 Sequence 84, Appl
35 37 36.6 251 9 US-09-895-913A-90 Sequence 90, Appli
36 37 36.6 331 10 US-09-881-752A-304 Sequence 304, App
37 37 36.6 341 10 US-09-815-915-5 Sequence 5, Appli
38 37 36.6 352 10 US-09-925-302-563 Sequence 563, App
39 37 36.6 377 10 US-09-954-697-15 Sequence 15, Appl
40 37 36.6 381 10 US-09-815-242-5199 Sequence 5199, Ap
41 37 36.6 438 9 US-09-959-845-6 Sequence 6, Appli
42 37 36.6 451 9 US-09-991-496-108 Sequence 108, App
43 37 36.6 451 10 US-09-874-923-108 Sequence 108, App
44 37 36.6 510 9 US-09-959-845-4 Sequence 4, Appli
45 37 36.6 549 9 US-09-959-845-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDFPNMSETDL 20
|||||
Db 931 DLEETAANSKDFPNMSETDL 950

RESULT 2
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950
|||||

RESULT 3
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950
|||||

RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950
|||||

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 931 DLEEIAANSKDFPNMSETDL 950
|||||

RESULT 6
US-09-864-761-41831
; Sequence 41831, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41831
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC025488.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EST_HUMAN HIT: BF689099.1, EVALUOE 7.00e-34
; OTHER INFORMATION: SWISSPROT HIT: P06133, EVALUOE 1.00e-37
US-09-864-761-41831

Query Match 40.6%; Score 41; DB 10; Length 72;
Best Local Similarity 56.7%; Pred. No. 3.5;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 AANSKDFPNMSETDL 20
||| ||| ||| |||
Db 47 AAVSLDFHMSSTD 61

RESULT 7

US-09-925-300-1440
; Sequence 1440, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1440
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1440

Query Match 40.6%; Score 41; DB 10; Length 459;
Best Local Similarity 33.3%; Pred. No. 32;
Matches 10; Conservative 3; Mismatches 7; Indels 10; Gaps 1;

QY 1 DLEEIAAN-----SKDFPNMSETDL 20
||| : | |||
Db 364 DMESVLRNLGWTDAFELGKADFSGMSQTDL 393

RESULT 8

US-09-810-997-1
; Sequence 1, Application US/09810997
; Patent No. US20020007501A1
; GENERAL INFORMATION:
; APPLICANT: Song, Xiaoling
; APPLICANT: Fan, Hao
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
; FILE REFERENCE: 21829/62
; CURRENT APPLICATION NUMBER: US/09/810,997
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/191,649
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/250,710
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-810-997-1

Query Match 39.6%; Score 40; DB 10; Length 284;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 EEIAANSKDFPNMSETDL 20
: | ||| ||| :
Db 174 QRIDVFSKDFDNIAEVEL 191

RESULT 9

US-09-815-242-11280
; Sequence 11280, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A

```
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11280
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-11280

Query Match      39.6%; Score 40; DB 10; Length 381;
Best Local Similarity 47.4%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DLEETAANSKDFPNMSETD 19
    ||||| : |||||
Db 41 DLENIALRQRVLKDMSELD 59

RESULT 10
US-09-864-761-35385
; Sequence 35385, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006657
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006654
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35385
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000406.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EST_HUMAN HIT: AW752708.1, EVALUE 1.00e-119
; OTHER INFORMATION: SWISSPROT HIT: P54654, EVALUE 2.40e-01
US-09-864-761-35385

Query Match      39.6%; Score 40; DB 10; Length 499;
Best Local Similarity 44.4%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 BEIAANSKDFPNMSETDL 20
    || : || |||
Db 482 BEVFVNQDDVMMSDEDL 499

RESULT 11
US-09-803-286A-10
; Sequence 10, Application US/09803286A
; Patent No. US20020062504A1
; GENERAL INFORMATION:
; APPLICANT: Tanksley, Steven D.
; APPLICANT: Brommonschenkel, Sergio H.
; TITLE OF INVENTION: PLANT GENE CONFERRING RESISTANCE TO TOSPOVIRUSES
; FILE REFERENCE: 19603/3201
; CURRENT APPLICATION NUMBER: US/09/803,286A
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/188,356
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Lycopersicon var.
US-09-803-286A-10

Query Match      38.6%; Score 39; DB 10; Length 331;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 DLEETAANSKDFPNMSETDL 20
    ||||| : ||| : |
Db 276 DLEEIPLSFADPTLEQIKL 295

RESULT 12
US-09-815-242-10657
```

; Sequence 10657, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10657
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10657

Query Match 38.6%; Score 39; DB 10; Length 341;
Best Local Similarity 42.1%; Pred. No. 48;
Matches 8; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 DLEET--AANSKDFPNMSE 17
||::|:||||:|
DB 225 DLSQVHEAIKTKDFPRLGE 243

RESULT 13
US-10-037-677-4
; Sequence 4, Application US/10037677
; Patent No. US20020173003A1
; GENERAL INFORMATION:
; APPLICANT: Schellenberger, Volker
; APPLICANT: Liu, Amy D.
; APPLICANT: Selifonova, Olga V.
; TITLE OF INVENTION: Directed Evolution of Microorganisms
; FILE REFERENCE: GC560
; CURRENT APPLICATION NUMBER: US/10/037,677
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/314,847
; PRIOR FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: E. blatte
US-10-037-677-4

Query Match 38.6%; Score 39; DB 9; Length 387;
Best Local Similarity 42.1%; Pred. No. 55;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 LEEIAANSKDFPNMSETDL 20
||::|:||||:|
DB 344 LRELGVKEADFPYNAEMAL 362

RESULT 14
US-09-815-242-10387
; Sequence 10387, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10387
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10387

Query Match 38.6%; Score 39; DB 10; Length 396;
Best Local Similarity 40.0%; Pred. No. 57;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
||::|:||||:|
DB 41 DLSEVALRQRILKNMSDLSL 60

RESULT 15
US-09-815-242-10492
; Sequence 10492, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078

```

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10492
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; S-09-815-242-10492

```

```

Query Match      38.6%; Score 39; DB 10; Length 417;
Best Local Similarity 75.0%; Pred. No. 61;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 8 NSKDFPNM 15
Db 116 NTKDFPNL 123

```

```

Search completed: February 3, 2003, 13:32:28
Job time : 7 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 ; Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEEIAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	1 GNNYHM	genome polyprotein
2	101	100.0	2227	1 GNNYHR	genome polyprotein
3	101	100.0	2227	1 GNNYTK	genome polyprotein
4	101	100.0	2227	1 GNNYHB	genome polyprotein
5	98	97.0	2230	1 GNNYSA	genome polyprotein
6	50	49.5	736	2 D90574	hypothetical prote
7	49	48.5	400	2 F88931	protein Rllgl1.1
8	45.5	45.0	930	2 A84668	Argonaute (AGO1)-l
9	45.5	45.0	1621	2 T15264	hypothetical prote
10	44	43.6	740	1 FOLJHD	gag polyprotein -
11	44	43.6	927	2 T38127	phosphoprotein - f
12	44	43.6	1046	2 S67786	hypothetical prote
13	43.5	43.1	754	1 BABOH	peptide-aspartate
14	43.5	43.1	1451	2 S65571	pattern formation
15	43.5	43.1	5107	2 T29144	partial CDS - Caen
16	43	42.6	152	2 A64610	hypothetical prote
17	43	42.6	152	2 A71904	hypothetical prote
18	43	42.6	231	2 H70407	probable di-trans,
19	43	42.6	285	2 B87426	rhodanese family p
20	43	42.6	454	2 T02100	hypothetical prote
21	43	42.6	598	2 H71336	probable cell divi
22	43	42.6	1072	2 A84112	alkaline amylopull
23	43	42.6	1201	2 T08603	kinesin-related pr
24	42.5	42.1	913	2 T15278	hypothetical prote
25	42.5	42.1	4717	2 T41581	hypothetical coile
26	42	41.6	289	2 T23342	hypothetical prote
27	42	41.6	365	2 C81050	cytochrome c oxida
28	42	41.6	365	2 F81826	probable cytochrom
29	42	41.6	538	2 G95015	ABC transporter, A

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with
A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20

|||||

Db 931 DLEEIAANSKDFPNMSETDL 950

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03903

R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDPNNSETDL 20
|||||
DB 931 DLEETAANSKDPNNSETDL 950

RESULT 3
GNVYK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1998 #sequence_revision 30-Jun-1998 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with other HAV strains
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDPNNSETDL 20
|||||
DB 931 DLEETAANSKDPNNSETDL 950

RESULT 4
GNVYH
genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
Vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhard
Virus Res. 8, 153-171, 1987
A:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP2>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein Vpg #status predicted <VPG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDPNNSETDL 20
|||||
DB 931 DLEETAANSKDPNNSETDL 950

RESULT 5
GNVYA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 2D; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tzarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
R:Tzarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure with other HAV strains
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL2>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypeptide; polypeptide
F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-495/Product: coat protein 1C #status predicted <C1C>
 F:496-795/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.0%; Score 98; DB 1; Length 2230;
 Best Local Similarity 95.0%; Pred. No. 6.6e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
 |||||..|||||..|||||..|||
 Db 935 DLEETAAANSKDFPNMSETDL 954

RESULT 6

D90574 hypothetical protein MYPU_5000 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: D90574
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
 A:Reference number: A99512; MUID:21267165; PMID:11353084
 A:Accession: D90574
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-736 <KUR>
 A:Cross-references: GB:AL415566; PID:g14089914; PIDN:CAC13673.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTIP
 C:Genetics:
 A:Gene: MYPU_5000
 A:Genetic code: SGC3

Query Match 49.5%; Score 50; DB 2; Length 736;
 Best Local Similarity 55.0%; Pred. No. 8.3;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
 ||..||..||..||..||..||
 Db 218 DLVSIYNKDAPEISEEDL 237

RESULT 7

F88931 protein RL1G11.1 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: F88931
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
 A:Accession: F88931
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-400 <STO>
 A:Cross-references: GB:chr_V; PIDN:AAC69076.1; PID:g2384851; GSPDB:GN00023; CESP:RL1G11.1
 A:Note: contains similarity to C4-type zinc fingers
 C:Genetics:
 A:Gene: RL1G11.1
 A:Map position: 5

Query Match 48.5%; Score 49; DB 2; Length 400;
 Best Local Similarity 47.4%; Pred. No. 5.9;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 2 LEEETAAANSKDFPNMSETDL 20
 ||..||..||..||..||..||
 Db 164 LENTSTNKKDFPHSSKEDV 182

RESULT 8

A84668 Argonaute (AGO1)-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: A84668
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: A84668
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-930 <STO>
 A:Cross-references: GB:AE002093; NID:g3885334; PIDN:AAC77862.1; GSPDB:GN00139
 C:Genetics: 1
 A:Gene: At2g27040
 A:Map position: 2
 C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 45.0%; Score 45.5; DB 2; Length 930;
 Best Local Similarity 42.3%; Pred. No. 57;
 Matches 11; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

QY 2 LEEETAAANSKDF-----PNMSET 18
 |||::| |||| |::|
 Db 153 LEEVSATSKDFVSRANGSPNGNES 178

RESULT 9

T15264 hypothetical protein F59E12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15264
 R:Johnson, D.
 Submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid F59E12.
 A:Reference number: Z18318
 A:Accession: T15264
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-1621 <JOH>

A:Cross-references: EMBL:AF003386; NID:g2088833; PID:g2088843; PIDN:AAB54259.1; GSPDB

A:Experimental source: strain Bristol N2; clone F59E12

C:Genetics:

A:Gene: CESP:F59E12.9

A:Map position: 2

A:Introns: 30/3; 55/1; 200/2; 299/2; 327/2; 369/3; 589/3; 860/1; 986/1; 1278/1; 1547/

Query Match 45.0%; Score 45.5; DB 2; Length 1621;
 Best Local Similarity 42.9%; Pred. No. 1.1e+02;
 Matches 9; Conservative 8; Mismatches 1; Indels 3; Gaps 1;

QY 2 LEEETAAANSKDF---PNMSETD 19
 :|| |::||| |::|
 Db 985 IKEIVASAKDFMPDPVSDSD 1005

RESULT 10

FOLJHD

gag polyprotein - squirrel monkey retrovirus SMRV-H

N:Contains: core protein p16; core protein p19; probable core protein p10; probable c
 C:Species: squirrel monkey retrovirus SMRV-H

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A31827

R:Oda, T.; Ikeda, S.; Watanabe, S.; Hatsushika, M.; Akiyama, K.; Mitsunobu, F.

Virology 167, 468-476, 1988

A:Title: Molecular cloning, complete nucleotide sequence, and gene structure of the protein

A:Reference number: A31827; MUID:89073750; PMID:3201749

A:Accession: A31827

A:Molecule type: DNA

A:Residues: 1-740 <ODA>

A:Cross-references: GB:M23385; NID:g332626; PIDN:AAA66451.1; PID:g807672

C:Genetics:

C:Superfamily: AIDS-related virus gag polyprotein

C:Keywords: core protein; polyprotein

F:1-163/Product: core protein p19 #status predicted <CP9>

F:164-318/Product: core protein p16 #status predicted <CP6>

F:319-648/Product: core protein p35 #status predicted <CP5>

F:649-740/Product: core protein p10 #status predicted <CP1>

Query Match 43.6%; Score 44; DB 1; Length 740;

Best Local Similarity 50.0%; Pred. No. 76;

Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

QY 1 DLEEIAA--NSKDFPNMSET 18

||||| |:|:|:

Db 214 DLEAAAYNPMPQTLNT 233

RESULT 11

T38127

phosphoprotein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Aug-2002

C:Accession: T38127; A57087; S54119

R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1997

A:Reference number: Z2172

A:Accession: T38127

A:Status: preliminary; translated from GB/EMBL/DBDJB

A:Molecule type: DNA

A:Residues: 1-927 <BAD>

A:Cross-references: EMBL:Z95334; PIDN:CAB08599.2; GSPDB:GNO0066; SPDB:SPAC20G8.05c

A:Experimental source: strain 972h-; cosmid C20G8

R:Fankhauser, C.; Raymond, A.; Cerutti, L.; Utag, S.; Hofmann, K.; Simanis, V.

Cell 82, 435-444, 1995

A:Title: The Schizosaccharomyces pombe cdc15 gene is a key element in the reorganization

A:Reference number: A57087; MUID:95360987; PMID:7634333

A:Accession: A57087

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 'MLTKSLQ', 28-927 <FAN>

A:Cross-references: GB:X86179

R:Fankhauser, C.; Raymond, A.; Cerutti, L.; Simanis, V.

submitted to the EMBL Data Library, April 1995

A:Description: The cdc15 gene is a key element in F-actin reorganization at mitosis.

A:Reference number: S54119

A:Accession: S54119

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 'MLTKSLQ', 28-870, 'CKFIKEL' <FA2>

A:Cross-references: EMBL:X86179

C:Genetics:

A:Gene: SPDB:SPAC20G8.05c; cdc15

A:Map position: 1

A:introns: 27/3; 58/2; 871/1

C:Superfamily: fission yeast scd2 protein; SH3 homology

C:Keywords: mitosis; phosphoprotein

F:873-924/Domain: SH3 homology <SH3>

Query Match 43.6%; Score 44; DB 2; Length 927;

Best Local Similarity 47.1%; Pred. No. 98;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 LELIANSKDFPNMSET 18
|:|:|:|:|:|:|
Db 69 LQELAASSADIPEVGS 85

RESULT 12
S67786
hypothetical protein YDL223c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D0843
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67786
R:Rasmussen, S.W.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S6778
A:Accession: S67786
A:Molecule type: DNA
A:Residues: 1-1046 <RAS>
A:A:Cross-references: EMBL:Z74271; NID:g1431374; PID:e253363; PID:g1431375; GSPDB:GN000
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YDL223C
A:Cross-references: SGD:S0002382
A:Map position: 4L

Query Match 43.6%; Score 44; DB 2; Length 1046;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 EETIANSKDFPNMSET 18
|:|:|:|:|:|:|
Db 733 QDIASDAKDFTNPNPET 748

RESULT 13
BAB0H
peptide-aspartate beta-dioxygenase (EC 1.14.11.16) - bovine
N:Alternate names: aspartyl (asparaginyl) beta-hydroxylase
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1993 #sequence_revision 10-Feb-1995 #text_change 11-Jun-1999
C:Accession: A42969; B39470; B39470; C39470; S27948
R:Jia, S.; VanDusen, W.J.; Diehl, R.E.; Kohl, N.E.; Dixon, R.A.; Elliston, K.O.; Ster
J. Biol. Chem. 267, 14322-14327, 1992
A:Title: cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-hydroxylase
A:Reference number: A42969; MUID:92332546; PMID:1378441
A:Accession: A42969
A:Molecule type: mRNA
A:Residues: 1-754 <JIA>
A:Cross-references: EMBL:M91213; NID:g162693; PIDN:AAA03563.1; PID:g162694
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBIP:108534)
R:Wang, Q.; VanDusen, W.J.; Petroski, C.J.; Garsky, V.M.; Stern, A.M.; Friedman, P.A.
J. Biol. Chem. 266, 14004-14010, 1991
A:Title: Bovine liver aspartyl beta-hydroxylase. Purification and characterization.
A:Reference number: A39470; MUID:91310689; PMID:1856229
A:Accession: A39470
A:Molecule type: protein
A:Residues: 289-328 <WAN>
A:Accession: B39470
A:Molecule type: protein
A:Residues: 615,'X','617-630,'XX','633-634,'X','636,'XX','639-641 <WA2>
A:Accession: C39470
A:Molecule type: protein
A:Residues: 311-347,'X',349,'X',351-373,'X',375-379,'X',381-382 <WA3>
C:Comment: This enzyme uses ferrous iron as a cofactor, and while beta-hydroxylating
C:Comment: Aspartic acid and asparagine residues in the EGF homology domain of certain
C:Superfamily: peptide-aspartate beta-dioxygenase; tetrairicopeptide repeat homology
C:Keywords: glycoprotein; oxidoreductase; transmembrane protein
F:2-56/Domain: intracellular #status predicted <INC>
F:57-78/Domain: transmembrane #status predicted <TRM>
F:289-754/Product: peptide-aspartate beta-dioxygenase, 56K form #status predicted <S6
F:311-754/Product: peptide-aspartate beta-dioxygenase, 52K form #status predicted <S2
F:337-370/Domain: tetrairicopeptide repeat homology <TTI>

;

F:371-404/Domain: tetra tricopeptide repeat homology <TT2>
F:13,96,466,702/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.1% Score 43.5; DB 1; Length 754;
Best Local Similarity 55.6% Pred. No. 93;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 DLEEIAANSKDFP-NMSE 17
| ||| : ||| | : |||
Db 82 DYEVLAKAKDFRYNLSE 99

RESULT 14

S65571
pattern formation protein GNOM - Arabidopsis thaliana
N:Alternate names: EMB30 protein
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 07-May-1999
C:Accession: S65571; S65572
C:Author: Busch, M.; Mayer, U.; Juergens, G.
C:Pl. Gen. Genet. 250, 681-691, 1996
C:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure
A:Reference number: S65571; MUID:96204508; PMID:8628228
A:Accession: S65571
A:Molecule type: DNA
A:Residues: 1-1451 <BUS>
A:Cross-references: EMBL:U36433; NID:g1209632; PID:g1209633
A:Accession: S65572
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-110, 'I', 112-866, 'G', 868-1451 <BUW>
A:Cross-references: EMBL:U36432; NID:g1209630; PID:g1209631
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: GNOM; EMB30
A:Introns: 246/3

Query Match 43.1% Score 43.5; DB 2; Length 1451;
Best Local Similarity 63.2% Pred. No. 2e+02;
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 DLEEIAANS-KDFPNMSET 18
|| |||| | ||| | |||
Db 1302 DLLEIAAGSQKDYRNMECT 1320

RESULT 15

T29144
partial CDS - Caenorhabditis elegans
A:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T29144
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid K11C4.
A:Reference number: Z20577
A:Accession: T29144
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5107 <PAU>
A:Cross-references: EMBL:U64854; PIDN:AAB18318.1; GSPDB:GN00023; CESP:unc-68
A:Experimental source: strain Bristol N2; clone K11C4
C:Genetics:
A:Gene: CESP:unc-68
A:Map position: 5
A:Introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/
/3; 3269/2; 3313/2; 3466/1; 3519/3; 3615/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Query Match 43.1% Score 43.5; DB 2; Length 5107;
Best Local Similarity 43.5% Pred. No. 8.4e+02;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 1 DLEEIAANS---KDFPNMSETDL 20
||| : | | : | ||| : |||
Db 3377 DLETVANNTMYSDVPNVYVDL 3399

Search completed: February 3, 2003, 13:15:40
Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:05:12 ; Search time 5.16667 Seconds
(without alignments)
160.554 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEETAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	100.0	2226	1	POLG_HPAV4
2	101	100.0	2226	1	POLG_HPAV8
3	101	100.0	2227	1	POLG_HPAVL
4	101	100.0	2227	1	POLG_HPAVL
5	101	100.0	2227	1	POLG_HPAVL
6	98	97.0	2230	1	POLG_HPAV5
7	97	96.0	2226	1	POLG_HPAV2
8	46	45.5	194	1	Y041_THEAC
9	44	43.6	740	1	GAG_SMRVH
10	44	43.6	927	1	CC15_SCHPO
11	43.5	43.1	754	1	ASPH_BOVIN
12	43.5	43.1	1451	1	EM30_ARATH
13	43	42.6	231	1	UPPS_AQUAE
14	43	42.6	350	1	40MT_COPJA
15	41.5	41.1	1378	1	RPOB_CAMJE
16	41	40.6	346	1	RFAP_HAEIN
17	41	40.6	376	1	PTI6_HUMAN
18	41	40.6	526	1	CLOS_CLOHI
19	41	40.6	528	1	UDBA_HUMAN
20	41	40.6	548	1	HLXB_VITBC
21	40	39.6	283	1	RUAP_SOYBN
22	40	39.6	331	1	PLSX_UREPA
23	40	39.6	381	1	LLDD_HAEIN
24	40	39.6	547	1	NLTP_HUMAN
25	40	39.6	824	1	NSFH_CAEEL
26	40	39.6	1087	1	EAL3_HUMAN
27	39.5	39.1	244	1	P29_MYCPN
28	39.5	39.1	521	1	VL2_HPVO4
29	39.5	39.1	552	1	YM81_MYCTU
30	39	38.6	221	1	YSCL_YERPE
31	39	38.6	223	1	YSCL_YERPE
32	39	38.6	253	1	PCRB_METJA
33	39	38.6	320	1	Y149_MYCPN

ALIGNMENTS

RESULT 1
POLG_HPAV4
ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43C).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP MEDLINE=91162758; PubMed=1705995;
RX Lenon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination".
RL J. Virol. 65:2056-2065(1991).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -I- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59809; AAA45469.1; --
DR MEROPS: C03.005; --
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.

P43668 salmonella
P33232 escherichia
P47049 saccharomyc
P49075 agaricus bi
Q01917 crithidia f
P53235 saccharomyc
P27542 chlamydia p
P18593 agrobacteri
P38315 saccharomyc
Q9hbm0 homo sapien
P27951 streptococc
Q9a199 a peptide m

34 39 38.6 394 1 UXUA_SALTY
35 39 38.6 396 1 LLDD_ECOLI
36 39 38.6 396 1 YJEB_YEAST
37 39 38.6 438 1 GUX3_AGABI
38 39 38.6 474 1 CC2H_CRIFA
39 39 38.6 642 1 YG22_YEAST
40 39 38.6 660 1 DNAK_CHLPN
41 39 38.6 673 1 VID3_AGRTP
42 39 38.6 674 1 YB66_YEAST
43 39 38.6 779 1 VEZA_HUMAN
44 39 38.6 1164 1 BAG_STRAG
45 38.5 356 1 MSAB_ACTAC

FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BF75 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
|||||
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 2

ID POLG_HP4V8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M59808; AAA45467.1; --
DR MEROPS; C03.005; --
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAAANSKDFPNMSETDL 20
|||||
DB 931 DLEETAAANSKDFPNMSETDL 950

RESULT 3

ID POLG_HP4VH STANDARD; PRT; 2227 AA.
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M14114; AAA45475.1; --
DR EMBL; M14707; AAA45465.1; --
DR EMBL; M14707; AAA45466.1; ALT_INIT.

```

DR EMBL; M16632; AAA45471.1; -.
DR PIR; A25981; GNNYHM.
DR PIR; A25914; GNNYHM.
DR PIR; A03905; A03905.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
FT VARIANT 2227 2227
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
DB 931 DLEEIAANSKDFPNMSETDL 950

RESULT 4
POLG_HPVLV STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dina D.;
RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----

DR EMBL; K02990; AAA45472.1; -.
DR PIR; A03903; GNNYHR.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
DB 931 DLEEIAANSKDFPNMSETDL 950

RESULT 5
POLG_HPVLV STANDARD; PRT; 2227 AA.
AC P13901; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
RA Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MBB).";
RL Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M20273; AAA45474.1; -
DR PIR; J50303; GNNYHB.
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDFPNMSETDL 20
|||||:|||||:|||||
DB 931 DLEETAANSKDFPNMSETDL 950

RESULT 6
POLG_HPAYS STANDARD; PRT: 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.

CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D00924; BAA00766.1; -
DR EMBL; X15461; CAA33490.1; -
DR PIR; A30470; GNNYSA.
DR PIR; S04885; S04885.
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR001205; RNA_pol_P3D.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 97.0%; Score 98; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 2.6e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETAANSKDFPNMSETDL 20
|||||:|||||:|||||
DB 935 DLEETAANSKDFPNMSETDL 954

RESULT 7
POLG_HPAV2 STANDARD; PRT: 2226 AA.
ID POLG_HPAV2
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.


```
CC15_SCHPO
ID CC15_SCHPO STANDARD; PRT; 927 AA.
AC Q09822; O14365;
DT 01-FEB-1996 (Rel. 33, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell division control protein 15.
GN CDC15 OR SPAC20G8.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=95360987; PubMed=7634333;
RA Fankhauser C., Reymond A., Cerutti L., Utzig S., Hofmann K.,
RA Simanis V.;
RT "The S. pombe cdc15 gene is a key element in the reorganization of F-
actin at mitosis."
RL Cell 82:435-444(1995).
RN [2]
REVIEWS TO N-TERMINUS.
RA Fankhauser C., Reymond A., Cerutti L., Utzig S., Hofmann K.,
RA Simanis V.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Horeby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: AFTER THE ONSET OF MITOSIS, FORMS A RING-LIKE STRUCTURE
WHICH CO-LOCALIZES WITH THE MEDIAL ACTIN RING. APPEARS TO MEDIATE
CYTOSKELETAL REARRANGEMENTS REQUIRED FOR CYTOKINESIS. ESSENTIAL
FOR VIABILITY.
CC -!- DEVELOPMENTAL STAGE: PEAKS IN EARLY MITOSIS BEFORE SEPTATION.
CC -!- DOMAIN: THE N-TERMINAL REGION IS IN A COILED COIL STRUCTURE.
CC -!- PTM: PHOSPHORYLATED
CC -!- SIMILARITY: CONTAINS 1 FCH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
-----
or send an email to license@isb-sib.ch).
CC EMBL; X86179; CAA60115.1; -.
DR EMBL; Z95334; CAB08599.2; -.
DR HSP; P07751; 1TUD.
DR InterPro; IPR001060; Cdc15_Fes_CIP4.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00611; FCH; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00055; FCH; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS0133; FCH; 1.
DR PROSITE; PS0002; SH3; 1.
DR Mitosis; Cytoskeleton; SH3 domain; Phosphorylation; Coiled coil.
KW DOMAIN 20 101 FCH.
FT DOMAIN 108 207 COILED COIL (POTENTIAL).
FT DOMAIN 866 927 SH3.
SQ SEQUENCE 927 AA; 102119 MW; FDCE7E0AAA3D247D CRC64;
Query Match 43.6%; Score 44; DB 1; Length 927;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 2 LEEIAANSKDFPNMSET 18
Db 69 LQELAASSADIEVGST 85
-----
RESULT 11
ASPH_BOVIN
ID ASPH_BOVIN STANDARD; PRT; 754 AA.
AC Q28056;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Aspartyl/asparaginyl beta-hydroxylase (EC 1.14.11.16) (Aspartate beta-
hydroxylase) (ASP beta-hydroxylase) (Peptide-aspartate beta-
dioxxygenase).
DE ASPH.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Liver, and Brain;
RX MEDLINE=92332546; PubMed=1378441;
RA Jia S., Vandusen W.J., Diehl R.E., Kohl N.E., Dixon R.A.F.,
RA Elliston K.O., Stern A.M., Friedman P.A.;
RT "cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-
hydroxylase."
RL J. Biol. Chem. 267:14322-14327(1992).
RN [2]
RP SEQUENCE OF 289-385 AND 615-641.
RC TISSUE=Liver;
RX MEDLINE=91310689; PubMed=1856229;
RA Wang Q., Vandusen W.J., Petroski C.J., Garsky V.M., Stern A.M.,
RA Friedman P.A.;
RT "Bovine liver aspartyl beta-hydroxylase. Purification and
characterization."
RL J. Biol. Chem. 266:14004-14010(1991).
CC -!- FUNCTION: SPECIFICALLY HYDROXYLATES AN ASP OR ASN RESIDUE IN
CERTAIN EPIDERMAL GROWTH FACTOR-LIKE (EGF) DOMAINS OF A NUMBER OF
PROTEINS.
CC -!- CATALYTIC ACTIVITY: Peptide L-aspartate + 2-oxoglutarate + O(2) =
peptide 3-hydroxy-L-aspartate + succinate + CO(2).
CC -!- COFACTOR: IRON.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
reticulum.
```

```

CC -1- PTM: MIGHT BE PROCESSED TO THE 56 kDa (AA 289-754) OR 52 kDa (AA
CC 311-754) FORMS IN THE LUMEN OF THE ENDOPLASMIC RETICULUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M91213; AAA03563.1; -
CC InterPro; IPR001440; TPR.
CC Oxidoreductase; Dioxxygenase; Iron; Transmembrane; Signal-anchor;
CC Endoplasmic reticulum.
CC DOMAIN 1 57
CC TRANSMEM 58 78
CC -----
CC DOMAIN 79 754
CC DOMAIN 9 12
CC DOMAIN 14 21
CC DOMAIN 318 328
CC CARBOHYD 96 96
CC CARBOHYD 466 466
CC CARBOHYD 702 702
CC SEQUENCE 754 AA; 84998 MW; 369593A1F0B558C8 CRC64;
CC -----
Query Match 43.1%; Score 43.5; DB 1; Length 754;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 DLEETIAANS-KDFP-NMSE 17
| 111:1111111111
DB 82 DYEVLAKANDFRNLS 99

RESULT 12
EM30_ARATH STANDARD; PRT; 1451 AA.
ID Q42510; Q38983;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pattern formation protein EMB30.
GN EMB30 OR GNOM OR ATL1G33980 OR F7A19.7 OR F16A14.20.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=cv. Wassilewskija, and cv. Columbia;
RX MEDLINE=94291195; PubMed=8020095;
RA Shevell D.E., Leu W.-M., Gillmor C.S., Xia G., Feldmann K.A.,
RA Chua N.-H.;
RT "EMB30 is essential for normal cell division, cell expansion, and
RT cell adhesion in Arabidopsis and encodes a protein that has
RT similarity to Sec7."
RL Cell 77:1051-1062(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia, and cv. CSR-1;
RX MEDLINE=96204508; PubMed=8628226;
RA Busch W., Mayer U., Juergens G.;
RT "Molecular analysis of the Arabidopsis pattern formation of gene
RT GNOM: gene structure and intragenic complementation.";
RL Mol. Gen. Genet. 250:681-691(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;

```

```

RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Dewar K.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maitl R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Unterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
CC -1- FUNCTION: MAY PERFORM A FUNCTION THAT AFFECTS CELL EXPANSION, THE
CC ORIENTATION OF THE PLANE OF CELL DIVISION, THE NUMBER OF CELL
CC DIVISIONS, AND CELL ADHESION THROUGHOUT PLANT DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: STEMS, LEAVES, FLOWERS, SILIQUES, FLORAL
CC INFLORESCENCE AND ROOTS.
CC -1- SIMILARITY: CONTAINS 1 SEC7 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U56140; AAB01205.1; -
CC EMBL; U56141; AAB01206.1; -
CC EMBL; U36432; AAA91150.1; -
CC EMBL; U36433; AAA91151.1; -
CC EMBL; AC007576; AAD39284.1; -
CC EMBL; AC068197; AAF79403.1; -
CC HSSP; Q99418; IPBV.
CC InterPro; IPR000904; Sec7.
CC Pfam; PF01369; Sec7; 1.
CC SMART; SM00222; Sec7; 1.
CC PROSITE; PS00190; SEC7; 1.
CC DOMAIN 557 752
CC MUTAGEN 658 658
CC CONFLICT 111 111
CC CONFLICT 867 867
CC SEQUENCE 1451 AA; 162618 MW; 666E21C74B426996 CRC64;
CC -----
Query Match 43.1%; Score 43.5; DB 1; Length 1451;
Best Local Similarity 63.2%; Pred. No. 70;
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 DLEETIAANS-KDFPNMSET 18
| 1111111111111111
DB 1302 DLLETAAGSQKDYRNMEGT 1320

RESULT 13
UPPS_AQUAE STANDARD; PRT; 231 AA.
ID UPPS_AQUAE
AC 067291;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP synthetase)
DE (Di-trans-poly-cis-decaprenylcistransferase) (Undecaprenyl diphosphate
DE synthase) (UDS).
GN UPPS OR AQ_1248.
OS Aquifex aeolicus.

```



```

CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC BETA' CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; ALI39075; CAB75116.1; -.
CC EMBL; X77304; CAA54509.1; -.
CC HSSP; Q9KWU7; 1HQM.
CC InterPro: IPR001572; RNA_pol_B.
CC Pfam: PF00562; RNA_pol_B; 1.
CC PROSITE; PS01166; RNA_POL_BETA; 1.
CC Transferrase; Transcription; DNA-directed RNA polymerase;
CC Complete proteome.
CC CONFLICT 338 347 NDLANGVDAA -> MTWLMALMQP (IN REF. 2).
CC CONFLICT 558 558 A -> R (IN REF. 2).
CC CONFLICT 671 671 C -> S (IN REF. 2).
CC CONFLICT 691 691 A -> R (IN REF. 2).
CC SEQUENCE 1378 AA; 155915 MW; AB7467C305028EB5 CRC64;

Query Match 41.1%; Score 41.5; DB 1; Length 1378;
Best Local Similarity 47.4%; Pred. No. 1.4e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

Oy 2 LEEITAANSKDFPNMSETDL 20
   :||| :||| :||| :|||
Db 858 IEEI---TKDIPNVKEDV 873

```

Search completed: February 3, 2003, 13:12:14
Job time : 7.16667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:07:47 ; Search time 21.5 Seconds
(without alignments)
191.672 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101

Sequence: 1 DLEETAAANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phase:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:
- 15: sp_rvirus:
- 16: sp_bacteriaph:
- 17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
2	101	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
3	101	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
4	101	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
5	101	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
6	101	100.0	251	12 Q9ENQ1	Q9enq1 hepatitis a
7	101	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
8	101	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
9	101	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
10	101	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
11	101	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
12	101	100.0	251	12 Q9ENN7	Q9enn7 hepatitis a
13	101	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
14	101	100.0	251	12 Q9ENN5	Q9enn5 hepatitis a
15	101	100.0	251	12 Q9ENN4	Q9enn4 hepatitis a
16	101	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a

17	101	100.0	1124	12	Q84780	hepatitis a
18	101	100.0	1161	12	Q05794	hepatitis a
19	101	100.0	2216	12	Q9WMA2	hepatitis a
20	101	100.0	2218	12	Q67824	hepatitis a
21	101	100.0	2225	12	Q9DL32	hepatitis a
22	101	100.0	2227	12	Q67825	hepatitis a
23	101	100.0	2227	12	Q67826	hepatitis a
24	101	100.0	2227	12	Q9WMA4	hepatitis a
25	101	100.0	2227	12	Q9WMA3	hepatitis a
26	101	100.0	2227	12	Q9WMA1	hepatitis a
27	101	100.0	2227	12	Q9WMA0	hepatitis a
28	101	100.0	2227	12	Q9IFH5	hepatitis a
29	101	100.0	2227	12	Q8VON6	hepatitis a
30	95	94.1	251	12	Q9ENQ8	hepatitis a
31	95	94.1	251	12	Q9ENQ7	hepatitis a
32	95	94.1	251	12	Q9ENQ3	hepatitis a
33	95	94.1	251	12	Q9ENQ0	hepatitis a
34	95	94.1	251	12	Q9ENP8	hepatitis a
35	95	94.1	251	12	Q9ENP6	hepatitis a
36	95	94.1	251	12	Q9ENP4	hepatitis a
37	95	94.1	251	12	Q9ENP3	hepatitis a
38	95	94.1	251	12	Q9ENP0	hepatitis a
39	95	94.1	251	12	Q9ENN9	hepatitis a
40	95	94.1	251	12	Q9ENN8	hepatitis a
41	95	94.1	251	12	Q9ENN3	hepatitis a
42	95	94.1	2218	12	Q67817	hepatitis a
43	95	94.1	2227	12	Q8QV03	hepatitis a
44	94	93.1	184	12	Q87092	simian hepa
45	91	90.1	251	12	Q9ENQ2	hepatitis a

ALIGNMENTS

RESULT 1

ID	Q9ENR1	PRELIMINARY;	PRT;	251 AA.
AC	Q9ENR1:			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DE	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)		
DE	Polyprotein (Fragment)			
OS	Hepatitis A virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Hepatovirus.			
OX	NCBI_TaxID=12092;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A1;			
RA	Fujiwara K.;			
RT	"hepatitis A virus.";			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB047652; BAB12160.1;			
FT	NON_TER	1		
FT	NON_TER	251		
SQ	SEQUENCE	251 AA;	28749 MW;	58A520D873893445 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	DLEETAAANSKDFPNMSETDL 20
DB	95	DLEETAAANSKDFPNMSETDL 114

RESULT 2

ID	Q9ENQ9	PRELIMINARY;	PRT;	251 AA.
AC	Q9ENQ9:			
DT	01-MAR-2001	(TrEMBLrel. 16, Created)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)		
DT	01-MAR-2001	(TrEMBLrel. 16, Last annotation update)		

```
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETAANSKDFPNMSETDL 20
|||||
95 DLEETAANSKDFPNMSETDL 114

Db

RESULT 3
Q9ENQ6
ID Q9ENQ6 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A162;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETAANSKDFPNMSETDL 20
|||||
95 DLEETAANSKDFPNMSETDL 114

Db

RESULT 4
Q9ENQ5
ID Q9ENQ5 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
```

```
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETAANSKDFPNMSETDL 20
|||||
95 DLEETAANSKDFPNMSETDL 114

Db

RESULT 5
Q9ENQ4
ID Q9ENQ4 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A201;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETAANSKDFPNMSETDL 20
|||||
95 DLEETAANSKDFPNMSETDL 114

Db

RESULT 6
Q9ENQ1
ID Q9ENQ1 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A206;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047662; BAB12170.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
```



```
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNNSETDL 20
    |||||
Db 95 DLEEIAANSKDFPNNSETDL 114

RESULT 7
Q9ENP9
ID Q9ENP9 PRELIMINARY; PRT; 251 AA.
AC Q9ENP9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A303;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047664; BAB12172.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CA5C1A CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNNSETDL 20
    |||||
Db 95 DLEEIAANSKDFPNNSETDL 114

RESULT 8
Q9ENP7
ID Q9ENP7 PRELIMINARY; PRT; 251 AA.
AC Q9ENP7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNNSETDL 20
    |||||
Db 95 DLEEIAANSKDFPNNSETDL 114

RESULT 9
Q9ENP5
```

```
ID Q9ENP5 PRELIMINARY; PRT; 251 AA.
AC Q9ENP5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNNSETDL 20
    |||||
Db 95 DLEEIAANSKDFPNNSETDL 114

RESULT 10
Q9ENP2
ID Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC Q9ENP2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A5;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIAANSKDFPNNSETDL 20
    |||||
Db 95 DLEEIAANSKDFPNNSETDL 114

RESULT 11
Q9ENP1
ID Q9ENP1 PRELIMINARY; PRT; 251 AA.
AC Q9ENP1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A503;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 95 DLEEIAANSKDFPNMSETDL 114
|||||

RESULT 12

Q9ENN7 PRELIMINARY; PRT; 251 AA.
AC Q9ENN7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A712;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047676; BAB12184.1; -
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28785 MW; BFE79D3A26134F18 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 95 DLEEIAANSKDFPNMSETDL 114
|||||

RESULT 13

Q9ENN6 PRELIMINARY; PRT; 251 AA.
AC Q9ENN6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A713;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -
FT NON_TER 1

FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 95 DLEEIAANSKDFPNMSETDL 114
|||||

RESULT 14

Q9ENN5 PRELIMINARY; PRT; 251 AA.
AC Q9ENN5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A75;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047678; BAB12186.1; -
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28658 MW; 98E8EED00B2EDF10 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
Db 95 DLEEIAANSKDFPNMSETDL 114
|||||

RESULT 15

Q9ENNA PRELIMINARY; PRT; 251 AA.
AC Q9ENNA;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A77;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEEIAANSKDFPNMSETDL 20
|||||

Db 95 DLEETIAANSKDFPNNSETDL 114

Search completed: February 3, 2003, 13:14:24
Job time : 22.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:04:52 ; Search time 27.8333 Seconds
(without alignments)
95.749 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRLGLSGVQBIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	20	AAW42929	Immunogenic Hepati
2	96	100.0	20	AAW42929	Synthetic HAV P2A
3	96	100.0	25	AAW42969	Immunogenic Hepati
4	96	100.0	25	AAW42969	Synthetic HAV P2A
5	96	100.0	2227	7 AAP60066	Sequence of viral
6	96	100.0	2227	11 AAR05697	Attenuated hepatitis
7	96	100.0	2227	18 AAW34074	Hepatitis A virus
8	96	100.0	2227	21 AAB18607	Amino acid sequenc
9	96	100.0	2227	21 AAB18608	Amino acid sequenc
10	96	100.0	2227	21 AAB18609	Amino acid sequenc

11	96	100.0	2227	23 AAE19899	Hepatitis A virus
12	43	44.8	80	22 AAU41911	Propionibacterium
13	43	44.8	738	22 ABP26213	Streptococcus poly
14	43	44.8	922	22 AAG85023	Shrimp white spot
15	42.5	44.3	410	23 ABP28012	Streptococcus poly
16	42	43.8	115	23 ABB55412	Lactococcus lactis
17	42	43.8	625	22 AAU49915	Propionibacterium
18	42	43.8	1035	22 AAB60277	Saccharomyces cere
19	41.5	43.2	2014	22 AAG67395	Amino acid sequenc
20	41.5	43.2	2014	22 AAE24137	Human kinase (PKIN
21	41	42.7	213	22 ABB59304	Drosophila melanog
22	41	42.7	291	22 ABG06041	Novel human diagno
23	41	42.7	332	20 AAY37809	Protein which is s
24	41	42.7	366	19 AAW60855	Mouse CD14 protein
25	41	42.7	366	21 AAY53879	A murine CD14 prot
26	41	42.7	605	22 AAU36494	Pseudomonas aerugi
27	41	42.7	708	19 AAW69844	Amino acid sequenc
28	41	42.7	708	23 ABB75006	Human clone L761P
29	41	42.7	715	22 ABG12254	Novel human diagno
30	41	42.7	748	19 AAW50143	Cyanobacterial phy
31	40	41.7	92	22 ABG19753	Novel human diagno
32	40	41.7	140	22 AAG90547	C glutamicum prote
33	40	41.7	144	22 AAU57103	Propionibacterium
34	40	41.7	144	22 AAU65829	Propionibacterium
35	40	41.7	154	22 ABB29032	Peptide #1683 enco
36	40	41.7	154	22 ABB34196	Peptide #1702 enco
37	40	41.7	154	22 ABB19633	Protein #1632 enco
38	40	41.7	154	22 AAM54987	Human brain expres
39	40	41.7	154	22 AAM67372	Human bone marrow
40	40	41.7	154	22 AAM15204	Peptide #1638 enco
41	40	41.7	154	22 AAM27663	Peptide #1700 enco
42	40	41.7	154	22 AAM02946	Peptide #1628 enco
43	40	41.7	154	23 ABB37000	Human peptide enco
44	40	41.7	315	22 AAG71918	Human olfactory re
45	40	41.7	315	22 AAG72722	Human olfactory re

ALIGNMENTS

RESULT 1
AAW42929
ID AAW42929 standard; peptide; 20 AA.
XX
AC AAW42929;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1331.
DE
DE Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
KW
XX Synthetic.
OS
OS Hepatitis A virus.
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fields HA, Khudyakov YE;
XX
XX WPI; 1997-535831/49.
XX
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal

XX PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 961-980, and has a reactivity of 27.1% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 96; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRLMLGLSGVQEIKEQ 20
Db 1 KINLADRLMLGLSGVQEIKEQ 20

RESULT 2
AAB69446
ID AAB69446 standard; Peptide; 20 AA.
AC AAB69446;
XX
XX 20-APR-2001 (first entry)
DE Synthetic HAV P2A peptide, SEQ ID NO: 46.
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
XX Hepatitis A virus.
OS Synthetic.
XX WO200105824-A2.
XX
XX 25-JAN-2001.
XX
XX 14-JUL-2000; 2000WO-US19267.
XX
XX 15-JUL-1999; 99US-0144412.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fields HA, Khudyakov YE;
XX WPI; 2001-112681/12.
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines -
XX
XX Claim 13; Page 97; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IgM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.

XX SQ Sequence 20 AA;
Query Match 100.0%; Score 96; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.7e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRLMLGLSGVQEIKEQ 20
Db 1 KINLADRLMLGLSGVQEIKEQ 20

RESULT 3
AAW42969
ID AAW42969 standard; peptide; 25 AA.
XX
AC AAW42969;
XX
XX 28-APR-1998 (first entry)
XX Immunogenic Hepatitis A virus peptide YK-1757.
XX
XX Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX Synthetic.
OS Hepatitis A virus.
XX WO9740147-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06891.
XX
XX 19-APR-1996; 96US-0015644.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fields HA, Khudyakov YE;
XX WPI; 1997-535831/49.
XX
XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
XX immune response to HAV in a mammal or to detect the presence of
XX antibodies against HAV in a mammal
XX
XX Claim 18; Page 112; 140pp; English.

XX The present immunogenic peptide corresponds to an immunogenic
XX epitope of the Hepatitis A virus (HAV). The peptide is substantially
XX similar to a portion of the amino acid sequence of the P2A protein of HAV
XX corresponding to amino acids 792-980. Compositions containing the
XX peptide can be used to induce an immune response to HAV in a mammal.
XX The peptide can also be used to detect the presence of antibodies
XX against HAV in mammalian serum. The peptide can also be used to make an
XX antibody against HAV by administering the peptide to a mammal.

XX SQ Sequence 25 AA;
Query Match 100.0%; Score 96; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRLMLGLSGVQEIKEQ 20
Db 6 KINLADRLMLGLSGVQEIKEQ 25

RESULT 4
AAB69448

ID AAB69448 standard; Peptide; 25 AA.
XX AAB69448;
AC
XX
DT 20-APR-2001 (first entry)
XX
XX Synthetic HAV P2A peptide, SEQ ID NO: 48.
DE
XX
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
XX Hepatitis A virus.
OS Synthetic.
OS
XX
XX WO200105824-A2.
PN
XX
XX 25-JAN-2001.
PD
XX
XX 14-JUL-2000; 2000WO-US19267.
PF
XX
XX 15-JUL-1999; 99US-0144412.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Fields HA, Khudyakov YE;
PI
XX
XX WPI; 2001-112681/12.
DR
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
PT
XX
XX Claim 13; Page 99; 130pp; English.
PS
XX
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
XX Sequence 25 AA;
SQ
Query Match 100.0%; Score 96; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. NO. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVQEIKEQ 20
DB 6 KINLADRMGLSGVQEIKEQ 25
RESULT 5
AAP60066
ID AAP60066 standard; Protein; 2227 AA.
XX
AC AAP60066;
XX
XX 26-JUN-1991 (first entry)
DT
XX
XX Sequence of viral I434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.
XX
XX Diagnosis; vaccine; passive immunotherapy.
KW

XX Hepatitis A virus.
OS
XX
PH Key Location/Qualifiers
XX 1..245
FT Region /label= P1.1A
FT Region 246..491
FT Region /label= 1B
FT Region 492..836
FT Region /label= 1C
FT Region 837..980
FT Region /label= P2.2A
FT Region 981..1076
FT Region /label= 2B
FT Region 1077..1422
FT Region /label= 2C
FT Region 1423..1484
FT Region /label= P3.3A
FT Region 1485..1507
FT Region /label= 3B
FT Region 1508..1678
FT Region /label= 3C
FT Region 1679..2227
FT Region /label= 3D
XX
XX EPI99480-A.
PN
XX
XX 29-OCT-1986.
PD
XX
XX 03-APR-1986; 86EP-0302465.
PF
XX
XX 03-APR-1985; 85US-0719329.
PR
XX
XX (CHIR-) CHIRON CORP.
PA
XX
XX Dina D, Potter SJ, Vannest GA, Caput D;
PI
XX
XX WPI: 1986-286213/44.
DR
XX
XX N-PSDB; AAN60080.
DR
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use
PT in prodn. of vaccines and diagnostic probes
PT
XX
XX Claim 5; Fig 1; 18pp; English.
PS
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
CC hepatitis A virus; transformed hosts may be used for expression of
CC polypeptides and fragments useful in vaccines without risk of
CC infection by the virus or in prodn. of particles which are capable
CC of inducing immunocompetent B cells for passive immunotherapy. Pref.
CC epitope is derived from AAs 445-657 or 792-848 of the HAV
CC polypeptide sequence (AAP60066).
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 96; DB 7; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 1.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980
RESULT 6
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX
XX
AC AAR05697;
XX
XX 15-AUG-1990 (first entry)
DT
XX
XX Attenuated hepatitis A virus.
DE

```

XX Hepatitis A virus; vaccine; attenuated.
KW Hepatitis A virus, strain HM-175.
OS
XX
XX
FH Key Location/Qualifiers
FT Region 1..23
FT /label=VP4 = 1A
FT Region 24..245
FT /label=VP2 = 1B
FT Region 246..491
FT /label=VP3 = 1C
FT Region 492..791
FT /label=VP1 = 1D
FT Region 792..980
FT /label=2A
FT Region 981..1087
FT /label=2B
FT Region 1088..1422
FT /label=2C
FT Region 1423..1496
FT /label=3A
FT Region 1497..1519
FT /label=3B = VPg
FT Region 1520..1738
FT /label=3C
FT Region 1739..2227
FT /label=3D
XX
US4894228-A.
PN
XX
XX
PD 16-JAN-1990.
XX
XX
PE 12-JUL-1988; 88US-0217824.
XX
PR 12-JUL-1988; 88US-0217824.
PR 12-JUL-1988; 88US-0652967.
XX
XX (USSH ) US DEPT HEALTH & HUMAN.
XX
XX Purcell RH, Ticehurst JR, Cohen I, Emerson SU, Feinstein SM;
PI Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
XX N-PSDB; AAQ03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel
PT attenuated hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 18pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine.
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 96; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980
RESULT 7
AAW34074
ID AAW34074 standard; Protein; 2227 AA.
XX

```

```

AC AAW34074;
XX
XX 27-APR-1998 (first entry)
XX
XX Hepatitis A virus HM-175 protein sequence.
DE
XX
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW infection; vaccine.
XX
XX Hepatitis A virus HM-175.
OS
XX
XX Key Location/Qualifiers
FH Protein 1..23
FT /label= VP4
FT Protein 24..245
FT /label= VP2
FT Protein 246..491
FT /label= VP3
FT Protein 492..791
FT /label= VP1
FT Protein 792..980
FT /label= 2A
FT Protein 981..1087
FT /label= 2B
FT Protein 1088..1422
FT /label= 2C
FT Protein 1423..1496
FT /label= 3A
FT Protein 1497..1519
FT /label= 3B
FT Protein 1520..1738
FT /label= 3C
FT Protein 1739..2227
FT /label= 3D
XX
XX WO9740166-A2.
PN
XX
XX 30-OCT-1997.
PD
XX
XX 18-APR-1997; 97WO-US06506.
XX
XX 19-APR-1996; 96US-0015642.
XX
XX (USSH ) US SEC DEPT HEALTH.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH, Raychaudhuri G;
PI
XX WPI; 1997-535850/49.
XX N-PSDB; AAT93023.
XX
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
PT as vaccines against HAV infection
XX
XX Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (1) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX

```


Query Match	100.0%;	Score 96;	DB 18;	Length 2227;
Best Local Similarity	100.0%;	Pred. No. 1.8e-07;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 KINLADRMGLSGVQEIKEQ 20			
DB	961 KINLADRMGLSGVQEIKEQ 980			
RESULT 8				
AAB18607				
ID	AAB18607 standard; Protein; 2227 AA.			
XX	AAB18607;			
XX				
DT	15-JAN-2001 (first entry)			
EE	Amino acid sequence of wild type Hepatitis A virus strain HM-175.			
XX				
KW	HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.			
XX				
OS	Hepatitis A virus.			
XX				
PN	US6113912-A.			
XX				
PD	05-SEP-2000.			
XX				
PF	07-JUN-1995; 95US-0475886.			
XX				
PR	18-SEP-1992; 92US-0947338.			
PR	17-SEP-1993; 93WO-US08610.			
PR	10-MAR-1995; 95US-0397232.			
XX				
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.			
XX				
PI	D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;			
XX				
DR	WPI: 2000-586464/55.			
DR	N-PSDB; AAA75476.			
XX				
PT	Novel live hepatitis A virus adapted to growth in human fibroblast cell			
PT	line useful as vaccine for protecting humans against hepatitis A virus			
PT	infection, has modified genome compared to wild type -			
XX				
PS	Disclosure; Fig 6A-K; 72pp; English.			
XX				
CC	The present sequence is derived from a wild type hepatitis A virus			
CC	(HAV) strain HM-174. The sequence is modified to produce HAV which			
CC	are adapted to growth in the human fibroblast-like cell line MRC-5.			
CC	The HAV is able to propagate in MRC-5 cells and retain appropriate			
CC	attenuation. It is useful as a live vaccine for prophylaxis of			
CC	hepatitis A in humans and other primates.			
XX				
XX				
SQ	Sequence 2227 AA;			
Query Match	100.0%;	Score 96;	DB 21;	Length 2227;
Best Local Similarity	100.0%;	Pred. No. 1.8e-07;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 KINLADRMGLSGVQEIKEQ 20			
DB	961 KINLADRMGLSGVQEIKEQ 980			
RESULT 9				
AAB18608				
ID	AAB18608 standard; Protein; 2227 AA.			
XX				
AC	AAB18608;			
XX				
DT	15-JAN-2001 (first entry)			

PR 10-MAR-1995; 95US-0397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
DR N-PSDB; AAA75478.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type -
XX
XX Disclosure; Columns 93-104; 72pp; English.
XX
XX The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.
XX
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 96; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980
RESULT 11
AAE19899
ID AAE19899 standard; Protein; 2227 AA.
XX
XX AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX Hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
OS
XX WO200213855-A2.
XX
XX 21-FEB-2002.
PD
XX
XX 15-AUG-2001; 2001WO-IB01808.
XX
XX 17-AUG-2000; 2000US-225767P.
PR 29-AUG-2000; 2000US-229175P.
PR 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPEP AB.
PA
XX
XX Sallberg M, Hultgren C;
PI
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus -
PT
XX Claim 11; Page 82-87; 120pp; English.
PS
XX The invention relates to a composition comprising ribavirin and an

CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to
CC a hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is hepatitis A virus (HAV) protein.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 96; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KINLADRMGLSGVQEIKEQ 20
Db 961 KINLADRMGLSGVQEIKEQ 980
RESULT 12
AAU41911
ID AAU41911 standard; Protein; 80 AA.
XX
XX AAU41911;
XX
XX 13-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #2807.
DE
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
OS
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
PA
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
DR N-PSDB; AAS59515.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 3106; 1069pp; English.
PS
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 80 AA;

Query Match 44.8%; Score 43; DB 22; Length 80;

Best Local Similarity 45.8%; Pred. No. 6.9;

Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 1;

QY 1 KINLADR---MLGLSGVQEIKEQ 20

DB 51 EIDVARRARCQCPNGVDDIAEQ 74

RESULT 13

ABP26213

ID ABP26213 standard; Protein; 738 AA.

XX AC ABP26213;

XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 1602.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 XX group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX OS Streptococcus agalactiae.

XX PN WO200234771-A2.

XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.

XX PR 27-OCT-2000; 2000GB-0026333.

XX PR 24-NOV-2000; 2000GB-0028727.

XX PR 17-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.

XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

XX PT Tettelin H;

XX DR WPI: 2002-352536/38.

XX DR N-PSDB; ABN66844.

XX PS New Streptococcus protein for the treatment or prevention of infection
 XX or disease caused by Streptococcus bacteria, such as meningitis, and
 XX for detecting a compound that binds to the protein -

XX Claim 1; Page 3311; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

XX Sequence 738 AA;

Query Match 44.8%; Score 43; DB 23; Length 738;

Best Local Similarity 36.8%; Pred. No. 91;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 KINLADRMLGLSGVQEIKE 19

DB 352 KVNQAEQSVGMGIQELVE 370

RESULT 14

AAG85023

ID AAG85023 standard; Protein; 922 AA.

XX AC AAG85023;

XX DT 11-SEP-2001 (first entry)

XX DE Shrimp white spot Bacilliform virus (WSBV) protein 114.

XX KW Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
 XX antiviral agent; gene expression; antisense construct;
 XX transgenic viral resistant shrimp.

XX OS White spot syndrome virus.

XX PN WO200138351-A2.

XX PD 31-MAY-2001.

XX PF 08-NOV-2000; 2000WO-US28888.

XX PR 24-NOV-1999; 99CN-0124717.

XX PA (PENY-) PE CORP NY.

XX PA (THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.

XX PA (SINO-) SINOGENOMAX CO LTD.

XX PI Xu X, Yang F, He J, Pham L, He M, Ye Y, Shen Y, Kodira C;

XX DR WPI: 2001-355877/37.

XX DR N-PSDB; AAG85023.

XX PT Primary nucleotide sequence of the shrimp white spot Bacilliform virus
 XX (WSBV), useful for producing viral polypeptides that can be used to
 XX screen for agents that are useful for treating WSBV infection -

XX Claim 1; Figure 3; 626pp; English.

XX The invention provides the primary nucleotide sequence of the WSBV genome
 CC (AAG82689), predicted transcript sequences (AAG62689-AAG62839) and
 CC encoded proteins (AAG84910-AAG85051) and oligonucleotide sequences
 CC (AAG62840-63160) suitable for use as primers or probes. The nucleic acid
 CC molecules and proteins of the invention are useful for diagnosis and
 CC monitoring viral infection, in screens for antiviral agents and for
 CC monitoring viral gene expression or activity during a treatment regimen.
 CC The nucleic acid molecules are also useful as antisense constructs to
 CC control viral gene expression in infected cells and tissues and to create
 CC transgenic viral resistant shrimp.

XX Sequence 922 AA;

Query Match 44.8%; Score 43; DB 22; Length 922;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 NLADRMGLSGVQEIKEQ 20

Search completed: February 3, 2003, 13:11:37
Job time : 29.8333 secs

Db 388 NLIRKTIGLSVEEKEE 405
|| : ||| | : | ||:
RESULT 15
ABP28012
ID ABP28012 standard; Protein: 410 AA.
XX AC
XX ABP28012;
XX DT
XX 02-JUL-2002 (first entry)
XX DE
XX Streptococcus polypeptide SEQ ID NO 5200.
XX KW
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS
XX Streptococcus agalactiae.
XX WO200234771-A2.
XX PN
XX
XX PD 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX PF
XX
XX 27-OCT-2000; 2000GB-0026333.
XX PR 24-NOV-2000; 2000GB-0028727.
XX PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
XX Tettelin H;
XX
XX WPI; 2002-352536/38.
XX DR N-PSDB; ABN68643.
XX
XX New Streptococcus protein for the treatment or prevention of infection
XX or disease caused by Streptococcus bacteria, such as meningitis, and
XX for detecting a compound that binds to the protein -
XX
XX Claim 1; Page 3863; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX antibodies that bind (I) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (I) are used to detect Streptococcus in a
XX biological sample. (I) is used to determine whether a compound binds to
XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (I) may be used to recombinantly produce (I) and may be
XX used in gene therapy. Antibodies to (I) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX
XX Sequence 410 AA;

Query Match 44.3%; Score 42.5; DB 23; Length 410;
Best Local Similarity 47.6%; Pred. No. 56;
Matches 10; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
QY 5 ADRM-----LGLSGVQEIKEQ 20
||| : | : ||| : |||
Db 98 ADRFGVIVASGIGGIQIEEQ 118

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:37 ; Search time 9.5 seconds
(without alignments)
61.943 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	2227	3	US-08-475-886-2
2	96	100.0	2227	3	US-08-475-886-4
3	96	100.0	2227	3	US-08-475-886-6
4	96	100.0	2227	4	US-08-397-232-2
5	96	100.0	2227	4	US-08-397-232-4
6	96	100.0	2227	4	US-09-171-387-2
7	96	100.0	2227	4	US-09-653-499-2
8	96	100.0	2227	4	US-09-653-499-4
9	96	100.0	2227	4	US-09-653-499-6
10	42	43.8	1250	1	US-08-441-139-9
11	41	42.7	366	3	US-08-746-883-6
12	41	42.7	708	4	US-09-643-597-369
13	40	41.7	332	1	US-08-469-649-2
14	40	41.7	332	4	US-09-347-878-60
15	40	41.7	638	2	US-08-846-762-95
16	40	41.7	643	4	US-09-178-252-25
17	40	41.7	1186	4	US-09-178-252-23
18	40	41.7	1207	1	US-07-951-715A-7
19	40	41.7	1207	2	US-08-459-448A-7
20	40	41.7	1207	3	US-08-459-595A-7
21	40	41.7	1207	3	US-08-459-504B-7
22	40	41.7	1207	3	US-08-459-444-7
23	40	41.7	1207	3	US-09-053-549-8
24	40	41.7	1207	4	US-09-547-422-7
25	40	41.7	1227	1	US-08-448-170-8
26	40	41.7	1227	3	US-09-053-549-2
27	40	41.7	1227	4	US-08-961-803-9

28	40	41.7	1229	1	US-08-100-709-4
29	40	41.7	1229	1	US-08-176-865-4
30	40	41.7	1229	1	US-08-474-038-4
31	40	41.7	1229	2	US-08-779-046-4
32	40	41.7	1229	2	US-08-881-340-4
33	40	41.7	1388	2	US-08-685-576-1
34	39	40.6	318	4	US-09-134-001C-4689
35	39	40.6	528	3	US-08-904-871-5
36	39	40.6	748	3	US-08-904-871-6
37	39	40.6	748	3	US-08-904-871-13
38	39	40.6	1211	4	US-09-134-001C-4820
39	38	39.6	185	2	US-08-691-814B-12
40	38	39.6	185	4	US-09-250-609-11
41	38	39.6	261	1	US-07-940-605A-2
42	38	39.6	261	1	US-08-184-422-8
43	38	39.6	261	1	US-08-360-923A-2
44	38	39.6	261	1	US-08-446-922-4
45	38	39.6	261	2	US-08-431-055-4

ALIGNMENTS

RESULT 1

US-08-475-886-2

; Sequence 2, Application US/08475886A

; Patent No. 6113912

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/08/475.886A

; CURRENT FILING DATE: 1995-06-07

; EARLIER APPLICATION NUMBER: 07/947,338

; EARLIER FILING DATE: 1992-09-18

; EARLIER APPLICATION NUMBER: 08/397,232

; EARLIER FILING DATE: 1995-03-10

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 2227

; TYPE: PRT

; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175

Query Match 100.0%; Score 96; DB 3; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20

|||||

Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 2

US-08-475-886-4

; Sequence 4, Application US/08475886A

; Patent No. 6113912

; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W

; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H

; APPLICANT: D'HONDT, ERIC

; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

; FILE REFERENCE: 20264262US2

; CURRENT APPLICATION NUMBER: US/08/475.886A

; CURRENT FILING DATE: 1995-06-07

; EARLIER APPLICATION NUMBER: 07/947,338

; EARLIER FILING DATE: 1992-09-18

; EARLIER APPLICATION NUMBER: 08/397,232

; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475.886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397.232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397.232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227

; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397.232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947.338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US97/06506
;; FILING DATE: 18-APR-1997
;; APPLICATION NUMBER: US60/015,642
;; FILING DATE: 19-APR-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: William S. Feller
;; REGISTRATION NUMBER: 26,728
;; REFERENCE/DOCKET NUMBER: 2026-4229US1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 758-4800
;; TELEFAX: (212) 751-6849
;; INFORMATION FOR SEQ ID NO: 2
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2227 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
|||||
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 7
US-09-653-499-2
;; Sequence 2, Application US/09653499
;; Patent No. 6423318
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT APPLICATION NUMBER: US/09/653,499
;; CURRENT FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
|||||
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 8
US-09-653-499-4
;; Sequence 4, Application US/09653499
;; Patent No. 6423318
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U

;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT APPLICATION NUMBER: US/09/653,499
;; CURRENT FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
|||||
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 9
US-09-653-499-6
;; Sequence 6, Application US/09653499
;; Patent No. 6423318
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT APPLICATION NUMBER: US/09/653,499
;; CURRENT FILING DATE: 2000-08-31
;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLGLSGVQEIKEQ 20
|||||
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 10
US-08-441-139-9
;; Sequence 9, Application US/08441139
;; Patent No. 5773245
;; GENERAL INFORMATION:
;; APPLICANT: Witttrup, Dr. Karl D.
;; APPLICANT: Robinson, Anne S.
;; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 863-4344
; TELEFAX: (416) 863-2653
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-746-883-6

Query Match 42.7% Score 41; DB 3; Length 366;
Best Local Similarity 63.6%; Pred. No. 15;
Matches 7; Conservative 4; Mismatches 0; Indels

QY 7 RMLGLSGVQEI 17
| : | : | : | : | :
Db 107 RVLGISLQL 117

RESULT 12
US-09-643-597-369
; Sequence 369, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643.597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 369
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-369

Query Match 42.7% Score 41; DB 4; Length 708;
Best Local Similarity 38.9%; Pred. No. 34;
Matches 7; Conservative 5; Mismatches 6; Indels

QY 3' NLADRMGLGSLGVQEIKEQ 20
| | : | : | : | : | :
Db 373 NHQKRSIGTGGIQDLKER 390

RESULT 13
US-08-469-649-2
; Sequence 2, Application US/08469649
; Patent No. 5728562
; GENERAL INFORMATION:
; APPLICANT: Shigyo, Tatsuro
; APPLICANT: Sugihara, Kohji
; APPLICANT: Takamoto, Yuji
; APPLICANT: Takashio, Masachika
; APPLICANT: Kamimura, Minoru
; APPLICANT: Yamamoto, Kazumi
; APPLICANT: Kojima, Yoshio
; APPLICANT: Kikuchi, Toshio
; APPLICANT: Emi, Shigenori
; TITLE OF INVENTION: AN ISOLATED URICASE

```


NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Frishauf, Holtz, Goodman, Langer &
ADDRESSEE: Chick, P.C.
STREET: 767 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10017-2023
COMPUTER READABLE FORM:
MEDIUM TYPE: 5, inch, 360 Kb Storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,649
FILING DATE: June 6, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Barth, Richard
REGISTRATION NUMBER: 28,180
REFERENCE/DOCKET NUMBER: 890578CIPD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 319-4900
TELEFAX: (212) 319-5101
TELEX: 236268
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 332 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-649-2

Query Match 41.7%; Score 40; DB 1; Length 332;
Best Local Similarity 41.2%; Pred. No. 21;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 INLADRLGLSGVQEI 18
Db 162 LNI TEQSSGLAGLQLIK 178
:|: :|:|:|:|:|

RESULT 14

US-09-347-878-60
Sequence 60, Application US/09347878C
Patent No. 6376210
GENERAL INFORMATION:
APPLICANT: Yuan, Chong
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES
FILE REFERENCE: 25885-1651
CURRENT APPLICATION NUMBER: US/09/347,878C
CURRENT FILING DATE: 1999-07-06
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 60
LENGTH: 332
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
OTHER INFORMATION: Urlicase protein sequence
US-09-347-878-60

Query Match 41.7%; Score 40; DB 4; Length 332;
Best Local Similarity 41.2%; Pred. No. 21;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 INLADRLGLSGVQEI 18
Db 162 LNI TEQSSGLAGLQLIK 178
:|: :|:|:|:|:|

RESULT 15

US-08-846-762-95
Sequence 95, Application US/08846762A
Patent No. 5994072
GENERAL INFORMATION:
APPLICANT: Lam, Joseph S.
APPLICANT: Burrows, Lori
APPLICANT: Charter, Deborah
APPLICANT: de Kievit, Teresa
TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assembly of O-Antigen in Pseudomonas Aeruginosa
FILE REFERENCE: 6580-089
CURRENT APPLICATION NUMBER: US/08/846,762A
CURRENT FILING DATE: 1997-04-30
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 95
LENGTH: 638
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-08-846-762-95

Query Match 41.7%; Score 40; DB 2; Length 638;
Best Local Similarity 47.4%; Pred. No. 46;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 INLADRLGLSGVQEI KEQ 20
Db 515 IDLAKRMNLSGLSIKSEE 533
:|:|:|:|:|:|:|:|:|:|

Search completed: February 3, 2003, 13:16:44
Job time : 10.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:14:27 : Search time 6 Seconds
(without alignments)
67.262 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRLMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	2227	9 US-10-104-966-12	Sequence 12, Appli
2	96	100.0	2227	9 US-10-135-988-2	Sequence 2, Appli
3	96	100.0	2227	9 US-10-135-988-4	Sequence 4, Appli
4	96	100.0	2227	9 US-10-135-988-6	Sequence 6, Appli
5	96	100.0	2227	10 US-09-929-955-12	Sequence 12, Appl
6	42	43.8	1250	10 US-09-801-368-364	Sequence 364, App
7	41	42.7	366	9 US-09-870-759-72	Sequence 72, Appl
8	41	42.7	705	10 US-09-815-242-12087	Sequence 12087, A
9	41	42.7	708	10 US-09-735-705-369	Sequence 369, App
10	41	42.7	708	10 US-09-850-716A-369	Sequence 369, App
11	41	42.7	1309	10 US-09-897-778-369	Sequence 369, App
12	41	42.7	1309	10 US-09-862-027-82	Sequence 82, Appl
13	40	41.7	140	9 US-09-738-626-4301	Sequence 4301, Ap
14	40	41.7	154	10 US-09-864-761-34931	Sequence 34931, A
15	40	41.7	549	9 US-09-815-242-10680	Sequence 10680, A
16	40	41.7	615	9 US-09-738-626-6067	Sequence 6067, Ap
17	40	41.7	643	10 US-09-826-660-25	Sequence 25, Appl
18	40	41.7	649	9 US-10-023-437-27	Sequence 27, Appl
19	40	41.7	1151	10 US-09-749-601A-10	Sequence 10, Appl

20	40	41.7	1186	10 US-09-826-660-23	Sequence 23, Appl
21	40	41.7	1770	10 US-09-801-368-298	Sequence 298, App
22	39	40.6	110	10 US-09-864-761-48195	Sequence 48195, A
23	39	40.6	365	10 US-09-815-242-11521	Sequence 11521, A
24	39	40.6	565	10 US-09-815-242-11867	Sequence 11867, A
25	39	40.6	748	10 US-09-272-809-1	Sequence 1, Appli
26	38.5	40.1	574	9 US-09-764-868-774	Sequence 774, App
27	38	39.6	185	10 US-09-250-611-11	Sequence 11, Appl
28	38	39.6	261	10 US-09-842-745A-2	Sequence 2, Appli
29	38	39.6	495	10 US-09-815-242-11215	Sequence 11215, A
30	37	38.5	184	10 US-09-250-611-13	Sequence 13, Appl
31	37	38.5	390	10 US-09-901-151-5	Sequence 5, Appli
32	37	38.5	390	10 US-09-901-151-6	Sequence 6, Appli
33	37	38.5	397	10 US-09-901-151-2	Sequence 2, Appli
34	37	38.5	414	10 US-09-901-151-4	Sequence 4, Appli
35	37	38.5	509	10 US-09-148-012-2	Sequence 2, Appli
36	37	38.5	838	9 US-09-712-363-150	Sequence 150, App
37	37	38.5	868	9 US-09-884-001-19	Sequence 19, Appl
38	36	37.5	91	9 US-09-738-626-6503	Sequence 6503, Ap
39	36	37.5	247	9 US-09-738-626-5773	Sequence 5773, Ap
40	36	37.5	338	10 US-09-815-242-5317	Sequence 5317, Ap
41	36	37.5	344	10 US-09-815-242-12624	Sequence 12624, A
42	36	37.5	416	8 US-08-900-220C-18	Sequence 18, Appl
43	36	37.5	416	10 US-09-151-999-18	Sequence 18, Appl
44	36	37.5	458	10 US-09-815-242-11759	Sequence 11759, A
45	36	37.5	491	10 US-09-801-368-78	Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 96; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGVQEIKEQ 20
Db 961 KINLADRLMLGLSGVQEIKEQ 980

RESULT 2
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U

; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 96; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 3
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 96; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 4
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988

; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 96; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 96; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980
|||||

RESULT 6
US-09-801-368-364
; Sequence 364, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John

```

; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801.368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 364
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-364

```

```

Query Match 43.8%; Score 42; DB 10; Length 1250;
Best Local Similarity 42.1%; Pred. No. 51;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 2 INLADRMGLSGVQEIKEQ 20
   :|: || || : ||| :
Db 871 LNILDRTLGEVVINEIKRK 889

```

```

RESULT 7
US-09-870-759-72
; Sequence 72, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-870-759-72

```

```

Query Match 42.7%; Score 41; DB 9; Length 366;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 7 RMLGLSGVQEI 17
   :||: ||: ||:
Db 107 RVLGISGLQEL 117

```

```

RESULT 8
US-09-815-242-12087
; Sequence 12087, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes

```

```

; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12087
; LENGTH: 605
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12087

```

```

Query Match 42.7%; Score 41; DB 10; Length 605;
Best Local Similarity 47.4%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy 1 KINLADRMGLSGVQEIKE 19
   | | | :|: ||| : ||
Db 18 KTTLVDKLLKLSGTLDRKE 36

```

```

RESULT 9
US-09-735-705-369
; Sequence 369, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 369
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-705-369

```

```

Query Match 42.7%; Score 41; DB 10; Length 708;
Best Local Similarity 38.9%; Pred. No. 39;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 3 NLADRMGLSGVQEIKEQ 20
   | | | :|: ||| : ||
Db 373 NHQKRISGTGGIQLDKER 390

```


APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34931
LENGTH: 154
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096770.14
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 9
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 6.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.7
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
OTHER INFORMATION: EST_HUMAN HIT: AI079550.1, EVALUATE 3.00e-24
OTHER INFORMATION: SWISSPROT HIT: O95371, EVALUATE 1.00e-31
US-09-864-761-34931

Query Match 41.7%; Score 40; DB 10; Length 154;
Best Local Similarity 63.6%; Pred. No. 10;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 8 MLGLSGVQEI 18
:|||||:

Db 13 LLGLTGVQELQ 23

RESULT 15

US-09-815-242-10680
; Sequence 10680, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.O11A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10680
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10680

Query Match 41.7%; Score 40; DB 10; Length 549;
Best Local Similarity 80.0%; Pred. No. 43;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KINLADRLMG 10
|||:|:
Db 527 KIRLADRLIG 536

Search completed: February 3, 2003, 13:32:29
Job time : 7 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:08:17 ; Search time 10.5 Seconds
(without alignments)
183.113 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRLMLGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	2227	1 GNNYHM	genome polyprotein
2	96	100.0	2227	1 GNNYHR	genome polyprotein
3	96	100.0	2227	1 GNNYMK	genome polyprotein
4	96	100.0	2227	1 GNNYHB	genome polyprotein
5	96	100.0	2230	1 GNNYSA	genome polyprotein
6	47	49.0	145	2 B25199	heat shock 16K pro
7	47	49.0	172	2 S33416	heat shock protein
8	44	45.8	414	2 T06303	enoyl-CoA hydratase
9	44	45.8	646	2 S72609	GTP-binding membra
10	44	45.8	653	2 G70683	probable lepA - My
11	44	45.8	788	2 S67595	hypothetical prote
12	43	44.8	145	2 B24289	heat shock protein
13	43	44.8	158	2 S64321	hypothetical prote
14	43	44.8	180	2 AC0056	conserved hypotnet
15	43	44.8	370	2 H70423	oxygen-independent
16	43	44.8	2104	2 D91286	hypothetical prote
17	43	44.8	2104	2 H86127	hypothetical prote
18	42.5	44.3	602	2 AB3542	gtp-binding protei
19	42	43.8	115	2 B86882	hypothetical prote
20	42	43.8	143	1 HHKW41	heat shock protein
21	42	43.8	547	2 A40656	hypothetical prote
22	42	43.8	547	2 AE1022	probable membrane
23	42	43.8	646	2 D81674	conserved hypotnet
24	42	43.8	1035	2 S18512	cell division cont
25	42	43.8	1250	2 A39578	SSD1 protein - yea
26	41	42.7	141	2 E84368	hypothetical prote
27	41	42.7	155	2 C34965	hypothetical 17K p
28	41	42.7	155	2 S15576	lipi protein - Shi
29	41	42.7	231	2 B98297	hypothetical prote

30	41	42.7	231	2 AE2986	conserved hypothet
31	41	42.7	313	2 D95866	probable transcrip
32	41	42.7	366	1 TDMSM4	monocyte surface g
33	41	42.7	518	1 S44183	thiamine-phosphate
34	41	42.7	605	2 D83007	regulatory protein
35	41	42.7	621	2 A71516	hypothetical prote
36	41	42.7	687	2 T09051	pepA protein - Pse
37	41	42.7	708	2 JC6329	yeast secretory pr
38	41	42.7	1036	2 F82263	probable multidrug
39	41	42.7	1338	2 T40993	protein kinase cek
40	41	42.7	1652	2 I50711	complement C3 prec
41	40	41.7	143	1 HHKW48	heat shock protein
42	40	41.7	208	2 A70122	glucose inhibited
43	40	41.7	247	2 F90656	hypothetical prote
44	40	41.7	247	2 F85507	hypothetical prote
45	40	41.7	332	2 JC4535	urate oxidase (EC

ALIGNMENTS

RESULT 1

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core

B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.8e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRLMLGLSGVQEIKEQ 20

Db 961 KINLADRLMLGLSGVQEIKEQ 980

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core

NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03903

R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||||
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 3

GNMYK

genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
R:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.H.; Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with other HAV sequences
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||||
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 4

GNMYH

genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
R:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.H.; Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with other HAV sequences
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
A:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhard
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP1>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
|||||
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 5

GNMYA

genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BAA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: JQ1080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-495/Product: coat protein 1C #status predicted <C1C>
 F:496-795/Product: coat protein 1D #status predicted <C1D>
 F:796-994/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2230;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 Db 965 KINLADRMGLSGVQEIKEQ 984

RESULT 6

heat shock 16K protein 2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 13-Aug-1999
 C:Accession: B25199
 R:Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.
 J. Biol. Chem. 261, 12006-12015, 1986

A:Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis elegans
 A:Reference number: A92555; MUID:86304344; PMID:3017958
 A:Accession: B25199
 A:Molecule type: DNA
 A:Residues: 1-145 <JON>
 A:Cross-references: GB:M14334; NID:G156338; PIDN:AAA28071.1; PID:G156340
 C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 145;
 Best Local Similarity 55.6%; Pred. No. 2.2;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKE 18
 |||||
 Db 65 KINLADRMGLSGVQEIKE 82

RESULT 7

S33416
 heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)
 C:Species: Nippostrongylus brasiliensis
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S33416

R:Tweddie, S.; Grigg, M.E.; Ingram, L.; Selkirk, M.E.
 submitted to the EMBL Data Library, April 1993
 A:Description: The expression of a small heat shock homologue is developmentally regulated in Nippostrongylus brasiliensis
 A:Reference number: S33416
 A:Accession: S33416
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-172 <TWE>
 A:Cross-references: EMBL:X71663; NID:G297865; PIDN:CAA50655.1; PID:G297866
 C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 172;
 Best Local Similarity 45.0%; Pred. No. 2.6;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKEQ 20
 |||||
 Db 80 KVLDDRLTVEGMEVETE 99

RESULT 8

T06303

enoyl-CoA hydratase homolog FlilC18.10 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
 C:Accession: T06303
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buyssehaert, C.; Dasseville, R.; De Clerck, R.;
 ewes, H.W.; Mayer, K.F.X.; Schueller, C.
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15589
 A:Accession: T06303
 A:Molecule type: DNA
 A:Residues: 1-414 <BEV>

A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:FlilC18.10
 A:Experimental source: cultivar Columbia; BAC clone FlilC18
 C:Genetics:
 A:Map position: 4
 A:Gene: ATSP:FlilC18.10
 A:Introns: 42/3; 75/2; 108/1; 134/3; 162/3; 191/1; 219/3; 260/2; 280/3; 313/3; 347/3;

Query Match 45.8%; Score 44; DB 2; Length 414;
 Best Local Similarity 52.6%; Pred. No. 22;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKE 19
 |||||
 Db 260 KIELIDKYFGLDTVEEIE 278

RESULT 9

S72609
 GTP-binding membrane protein lepA - Mycobacterium leprae
 N:Alternate names: protein B1937_f3_81
 C:Species: Mycobacterium leprae

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
 C:Accession: S72609
 R:Smith, D.R.; Robinson, K.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S72580
 A:Accession: S72609
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-646 <SMI>
 A:Cross-references: EMBL:U00016; NID:G466961; PIDN:AAA17177.1; PID:G466991
 C:Genetics:
 A:Gene: lepA
 A:Start codon: GTG
 C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu ho

C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
 F:45-177/Domain: translation elongation factor Tu homology <ETU>
 F:51-58/Region: nucleotide-binding motif A (P-loop)
 F:174-177/Region: GTP-binding NKXD motif

Query Match 45.8%; Score 44; DB 2; Length 646;
 Best Local Similarity 71.4%; Pred. No. 35;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGV 14
 |||||
 Db 57 KSTLADRMQLTGV 70

RESULT 10

G70683

probable lepA - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Feb-2001
 C:Accession: G70683

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: G70683
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-653 <COL>
A:Cross-references: GB:Z81368; GB:AL123456; NID:g3261656; PIDN:CAB03723.1; PID:g1655655
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: lepA
C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homolog
C:Keywords: GTP binding; nucleotide binding; P-loop
F:53-181/Domain: translation elongation factor Tu homology <ETU>
F:59-66/Region: nucleotide-binding motif A (P-loop)
F:178-181/Region: GTP-binding NXXD motif

Query Match 45.8%; Score 44; DB 2; Length 653;
Best Local Similarity 71.4%; Pred. No. 36;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGV 14
I | | | | | | | | | |
Db 65 KSTLADRLMLQLTGV 78

RESULT 11
S67595
hypothetical protein YDL060w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D2544
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67595
R:Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67587
A:Accession: S67595
A:Molecule type: DNA
A:Residues: 1-788 <BLO>
A:Cross-references: EMBL:Z74108; NID:g1431062; PID:g1431063; GSPDB:GN00004; MIPS:YDL060w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:TSR1; MIPS:YDL060w
A:Cross-references: SGD:S0002218
A:Map position: 4L

Query Match 45.8%; Score 44; DB 2; Length 788;
Best Local Similarity 61.5%; Pred. No. 44;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 MLGLSGVQEIKEQ 20
: | | | | | | | | | |
159 VFGLSGVQEVDE 171

RESULT 12
B24289
heat shock protein 16-1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 28-Sep-1987 #sequence_revision 30-Sep-1993 #text_change 29-Oct-1999
C:Accession: B24289; S31037; T25927; T25930
R:Rusnak, R.H.; Candido, E.P.M.
Mol. Cell. Biol. 5, 1268-1278, 1985
A:Reference number: A24289; MUID:85295957; PMID:4033652
A:Accession: B24289
A:Molecule type: DNA
A:Residues: 1-145 <RUS>
A:Cross-references: EMBL:K03273; NID:g156333; PIDN:AAA28068.1; PID:g156336
A:Note: the author translated the codon CAT for residue 17 as Tyr
R:Kay, R.J.; Rusnak, R.H.; Jones, D.; Mathias, C.; Candido, E.P.M.
Nucleic Acids Res. 15, 3723-3741, 1987
A:Title: Expression of intron-containing C. elegans heat shock genes in mouse cells demonstrates effect of heat shock on the mammalian splicing apparatus.
A:Reference number: S31036; MUID:87231065; PMID:3588308
A:Accession: S31037
A:Status: translation not shown

A:Molecule type: DNA
A:Residues: 41-63 <KAY>
A:Cross-references: EMBL:M31340; NID:g156343; PIDN:AAA28073.1; PID:g552068
R:Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid T27E4.
A:Reference number: Z20111
A:Accession: T25927
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-145 <BRA>
A:Cross-references: EMBL:U64837; PIDN:AAB04839.1; GSPDB:GN000023; CESP:hsp-16A
A:Experimental source: strain Bristol N2; clone T27E4
A:Accession: T25930
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-145 <BR2>
A:Cross-references: EMBL:U64837; PIDN:AAB04842.1; GSPDB:GN000023; CESP:hsp-16A
A:Experimental source: strain Bristol N2; clone T27E4
C:Genetics:
A:Gene: CESP:hsp-16A
A:Map position: 5
A:Introns: 42/3
C:Superfamily: alpha-crystallin

Query Match 44.8%; Score 43; DB 2; Length 145;
Best Local Similarity 45.0%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGVQEIKEQ 20
I | | | | | | | | | |
Db 65 KINLGDHTLSIQEGELKTE 84

RESULT 13
S64321
hypothetical protein YGR030c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein G4068
C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C:Accession: S64321
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64321
A:Molecule type: DNA
A:Residues: 1-158 <RIE>
A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN000
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:POP6; MIPS:YGR030c
A:Cross-references: SGD:S0003262
A:Map position: 7R

Query Match 44.8%; Score 43; DB 2; Length 158;
Best Local Similarity 64.3%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KINLADRLMLGLSGV 14
: | | | | | | | | | |
Db 65 QINMADRSILGLQGV 78

RESULT 14
AC0056
conserved hypothetical protein YP00454 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AC0056
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AC0056
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-180 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89310.1; PID:g15978546; GSPDB:GN00175
 A:Gene: YPO0454
 C:Superfamily: *Escherichia coli* conserved yjx protein

Query Match 44.8%; Score 43; DB 2; Length 180;
 Best Local Similarity 47.1%; Pred. No. 13;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 LADRMGLSGVQEIKEQ 20
 | | | | | : | : |
 Db 120 LGDEMAVLSCISNVKQ 136

RESULT 15
 .70423

oxygen-independent coproporphyrinogen III oxidase - *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
 C:Accession: H70423
 R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: H70423

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-370 <NOF>

A:Cross-references: GB:AE000739; NID:g2983813; PIDN:AAC07371.1; PID:g2983815; GB:AE00065

A:Experimental source: strain VFS

C:Genetics:

A:Gene: hemF

C:Superfamily: oxygen-independent coproporphyrinogen oxidase

Query Match 44.8%; Score 43; DB 2; Length 370;
 Best Local Similarity 66.7%; Pred. No. 28;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 6 DRMLGSGVQEI 17
 | : | | | | : | : |
 Db 78 DKVFGLSGVKEI 89

Search completed: February 3, 2003, 13:15:42
 Job time : 12.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:05:12 : Search time 5.16667 Seconds
(without alignments)
160.554 Million cell updates/sec

Title: US-09-171-432a-46
Perfect score: 96
Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2226	1	P26580 hepatitis a
2	96	100.0	2226	1	P26581 hepatitis a
3	96	100.0	2226	1	P26582 hepatitis a
4	96	100.0	2227	1	P08617 hepatitis a
5	96	100.0	2227	1	P06441 hepatitis a
6	96	100.0	2227	1	P13901 hepatitis a
7	96	100.0	2230	1	P14553 simian hepa
8	48	50.0	622	1	LEPA_STRCO
9	47	49.0	145	1	HS12_CAREL
10	47	49.0	172	1	HS20_NIPBR
11	44	45.8	646	1	LEPA_MYCLE
12	44	45.8	653	1	LEPA_MYCTU
13	43	44.8	145	1	HS11_CAREL
14	43	44.8	158	1	POP6_YEAST
15	42	43.8	143	1	HS16_CAREL
16	42	43.8	151	1	SODC_HAIRO
17	42	43.8	547	1	YJDB_SALTY
18	42	43.8	621	1	Y708_CHLMU
19	42	43.8	1035	1	CC68_YEAST
20	42	43.8	1250	1	SSD1_YEAST
21	41	42.7	141	1	NIKR_MENJA
22	41	42.7	155	1	TPPI_SHIFL
23	41	42.7	313	1	CBRL_RHIME
24	41	42.7	313	1	CBR2_RHIME
25	41	42.7	366	1	CD14_MOUSE
26	41	42.7	518	1	TH14_SCHPO
27	41	42.7	621	1	Y425_CHLTR
28	41	42.7	1338	1	CEK1_SCHPO
29	40	41.7	143	1	HS17_CAREL
30	40	41.7	208	1	GIDB_BORBU
31	40	41.7	316	1	OCDD3_HUMAN
32	40	41.7	502	1	URIC_BACSB
33	40	41.7	698	1	MCHF_ECOLI

34	40	41.7	739	1	RELA_STREQO	054089 streptococc
35	40	41.7	819	1	ION_CHLPN	0929f4 chlamydia p
36	40	41.7	1042	1	SYI_BORBU	051773 borrelia bu
37	40	41.7	1227	1	C1BE_BACTU	085805 bacillus th
38	40	41.7	1228	1	C1BA_BACTK	045517 bacillus th
39	40	41.7	1229	1	C1BB_BACTU	045739 bacillus th
40	40	41.7	1231	1	C1BC_BACTM	092a25 bacillus th
41	40	41.7	1233	1	C1BD_BACTM	045774 bacillus th
42	40	41.7	1770	1	R115_YEAST	P43565 saccharomyc
43	39	40.6	119	1	Y033_ARCFU	030238 archaeoglob
44	39	40.6	251	1	ASTA_ASTFL	P07584 astacus flu
45	39	40.6	365	1	AROC_HELPJ	0921h1 helicobacte

ALIGNMENTS

RESULT 1	ID	POLG_HPAV2	STANDARD:	PRT:	2226 AA.
AC	P26580:				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Genome polypeptide [contains: Coat proteins VP1 to VP4; Core proteins P2A to P2C; Probable proteins P3A to P3C; RNA-directed RNA polymerase P3D (PC 2.7.7.48)].				
DE	P3D (PC 2.7.7.48)].				
OS	Hepatitis A virus (strain 24a).				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.				
OX	NCBI_TaxID=12094;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=91162758; PubMed=1705995;				
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,				
RA	Cromeans T., Jansen R.W.;				
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination."				
RL	J. Virol. 65:2056-2065(1991).				
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).				
CC	-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.				
CC	-1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.				
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; M59810; AAA45468.1; -.				
DR	MEROPS; C03.005; -.				
DR	InterPro; IPR000605; RNA_helicase.				
DR	InterPro; IPR001205; RNA_pol_P3D.				
DR	Pfam; PF00680; RNA_dep_RNA_pol_1.				
DR	Pfam; PF00910; RNA_helicase_1.				
KW	Polypeptide; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.				
FT	CHAIN	1	23		
FT	CHAIN	24	245		
FT	CHAIN	246	791		
FT	CHAIN	492	794		
FT	CHAIN	795	900		
FT	CHAIN	901	1087		
FT	CHAIN	1088	1422		
FT	CHAIN	1423	1495		
FT	CHAIN	1496	1518		

FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVGEIKQ 20
Db 961 KINLADRMGLSGVGEIKQ 980
|||||

RESULT 2
POLG_HPVA4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RT J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59809: AAA5469.1; -
CC DR MEROPS: C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403BA08B09BF75 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVGEIKQ 20
Db 961 KINLADRMGLSGVGEIKQ 980
|||||

RESULT 3
POLG_HPVA8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RT J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59808: AAA5467.1; -
CC DR MEROPS: C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A6396CBDBB CRC64;

Oy 1 KINLADRLGLSGVQEIKEQ 20
 Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 4
 POLG_HPAVL STANDARD: PRT: 2227 AA.

AC P06617: P06443: 081082;
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]
 DE Hepatitis A virus (strain HM-175).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 NCBI_Taxid=12098;
 [1]
 AN SEQUENCE FROM N.A.
 AP STRAIN=Wild type;
 RC MEDLINE=87061253; PubMed=3023706;
 RX Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.;
 RA "Complete nucleotide sequence of wild-type hepatitis A virus: RT comparison with different strains of hepatitis A virus and other RT Picornaviruses";
 RT J. Virol. 61:50-59(1987).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Attenuated;
 RC MEDLINE=87175701; PubMed=3031686;
 RX Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M., Purcell R.H.;
 RA "Complete nucleotide sequence of an attenuated hepatitis A virus: RT comparison with wild-type virus";
 RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 RL [3]
 RN SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RP MEDLINE=85166289; PubMed=2984684;
 RX Baroudy B.M., Ticehurst J.R., Mele T.A., Malzel J.V. Jr., Purcell R.H., Feinstein S.M.;
 RA "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 RL [1]
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: STRAIN HM-175/77 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14107; AAA45475.1; -;
 DR EMBL: M14707; AAA45465.1; -;
 DR EMBL: M16632; AAA45466.1; ALT_INIT.
 DR EMBL: M16632; AAA45471.1; -;
 DR PIR: A25981; GNNYHM.
 DR PIR: A25914; GNNYMK.
 DR PIR: A03905; A03905.

DR MEROPS: C03.005; -;
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol: 1.
 DR Pfam: PF00910; RNA_helicase: 1.
 DR Pfam: PF00910; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1052 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE740A6 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. NO. 1.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20
 Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 5
 POLG_HPAVL STANDARD: PRT: 2227 AA.
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]
 DE Hepatitis A virus (strain LA).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 NCBI_Taxid=12099;
 [1]
 AN SEQUENCE FROM N.A.
 AP MEDLINE=85190549; PubMed=2986127;
 RX Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., van Nest G., Dina D.;
 RA "Primary structure and gene organization of human hepatitis A virus";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 RL [1]
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02990; AAA5472.1; -.
DR PIR: A03903; GNNTHR.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
DR Pfam: PF00910; RNA_helicase. 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIKED 20
Db 961 KINLADRMGLSGVGEIKED 980

RESULT 6
POLG_HPAVM STANDARD; PRT; 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HMB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wissel T., Klehn R., Wimmer E.,
RA Delnhardt F.;
RT "The encode nucleotide sequence of the genome of human hepatitis A
RT virus (isolate HMB).";
RT Virus Res. 8:153-171(1987).
RL
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN C3.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20273; AAA5474.1; -.
DR PIR: J50303; GNNTHR.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001205; RNA_pol_P3D.
DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
DR Pfam: PF00910; RNA_helicase. 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; BC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIKED 20
Db 961 KINLADRMGLSGVGEIKED 980

RESULT 7
POLG_HPAVS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains.";
RT J. Gen. Virol. 72:1677-1683(1991).
RL
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D00924; BAA00766.1; -;
 DR EMBL: X15461; CA33490.1; -;
 DR PIR: A30470; GNYSR.
 DR PIR: S04885; S04885.
 DR MEROPS: C03.005; -;
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR Pfam: PF00679; EFG_C; 1.
 DR Pfam: PF03144; GTP_EFTU_D2; 1.
 DR PRINTS: PR00315; ELONGATNCT.
 DR TIGRFAMs: TIGR00231; small_GTP; 1.
 DR PROSITE: PS00301; EFACOR.GTP; 1.
 KW NP_BIND; Complete proteome.
 FT NP_BIND 26 33 GTP (BY SIMILARITY).
 FT NP_BIND 94 98 GTP (BY SIMILARITY).
 FT NP_BIND 148 151 GTP (BY SIMILARITY).
 SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2230;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVOEIKQ 20
 Db 965 KINLADRLGLSGVOEIKQ 984

RESULT 8
 LEPA_STRCO STANDARD; PRT; 622 AA.

ID LEPA_STRCO
 AC 09RDC9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GTP-binding protein lepa.
 GN LEPA OR SC02562 OR SCC77.29C.

OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID:1902;

[1]

RP SEQUENCE FROM N.A.
 RC SRRAIN-A3(2) / M145;
 RX MEDLINE-21996410; PubMed-12000953;

RA Bentley S.D., Chater K.F., Gerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabbittowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrett B.G., Parkhill J.,
 RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2)".
 RL Nature 417:141-147(2002).
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL136503; CAB66240.1; -;
 DR HSPR: P13551; 2EFG.
 DR InterPro: IPR000640; EFG_C.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF00679; EFG_C; 1.
 DR Pfam: PF03144; GTP_EFTU_D2; 1.
 DR PRINTS: PR00315; ELONGATNCT.
 DR TIGRFAMs: TIGR00231; small_GTP; 1.
 DR PROSITE: PS00301; EFACOR.GTP; 1.
 KW NP_BIND; Complete proteome.
 FT NP_BIND 26 33 GTP (BY SIMILARITY).
 FT NP_BIND 94 98 GTP (BY SIMILARITY).
 FT NP_BIND 148 151 GTP (BY SIMILARITY).
 SQ SEQUENCE 622 AA; 68378 MW; 83F5C76FA2A80C7C CRC64;

Query Match 50.0%; Score 48; DB 1; Length 622;
 Best Local Similarity 57.9%; Pred. No. 3.3;
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVOEIKQ 19
 Db 32 KSTLADRLGLTGVVEQRO 50

RESULT 9
 HS12_CAEEL STANDARD; PRT; 145 AA.

ID HS12_CAEEL
 AC P06582;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein HSP16-2.

GN HSP16-2 OR Y46H3A.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID:6239;

[1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-86304344; PubMed-2017958;
 RA Jones D., Russnak R.H., Kay R.J., Candido E.P.M.;

RT "Structure, expression, and evolution of a heat shock gene locus in
 RT Caenorhabditis elegans that is flanked by repetitive elements.";

RL J. Biol. Chem. 261:12006-12015(1986).

[2]

RP SEQUENCE FROM N.A.
 RC SRRAIN-Bristol N2;
 RA Stoneking T., Wohlmann P., Lennox S.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

[3]

RP REVISIONS.

RA Watson R.;

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP-binding protein lepa.
 GN LEPA OR RV2404C OR MT2476 OR MTCY253.16.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid:1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37Rv;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Helt D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ertolaeva M.D., Salzberg S.L.,
 RA Kachner A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: Z81368; CAB03723.1; -;
 DR EMBL: AE007086; AAK46772.1; -;
 DR HSSP: P13551; ZEEG.
 DR TIGR: MT2476; -;
 DR TuberculList: RV2404C; -;
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR000795; EF-GTPbind.
 DR InterPro: IPR005225; Small-GTP.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR Pfam: PF03144; GTP_EFTU_D2; 1.
 DR TIGRFAMs: TIGR00231; small_GTP; 1.
 DR PROSITE: PS00301; EFACITOR_GTP; 1.
 KM GTP-binding; Complete Proteome.
 FT NP_BIND 59 66 GTP (BY SIMILARITY).
 FT NP_BIND 124 128 GTP (BY SIMILARITY).
 FT NP_BIND 178 181 GTP (BY SIMILARITY).
 SO SEQUENCE 653 AA; 72395 MW; DA4AFE10E6C5755 CRC64;

Query Match 45.88; Score 44; DB 1; Length 653;
 Best Local Similarity 71.48; Pred. No. 16;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGV 14
 Db 65 KSTLADRLGLTGV 78

RESULT 13

HS11_CABEL
 ID HS11_CABEL STANDARD; PRT; 145 AA.
 AC P34696;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein HSP16-1.
 GN (HSP16-1A OR T27E4.2) AND (HSP16-1B OR T27E4.8).
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_Taxid:6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85295957; PubMed-4033652;
 RA Russnak R.H., Candido E.P.M.;
 RT "Locus encoding a family of small heat shock genes in Caenorhabditis
 RT elegans: two genes duplicated to form a 3.8-kilobase inverted
 RT repeat."
 RL Mol. Cell. Biol. 5:1268-1278(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Bradshaw H.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 41-63 FROM N.A.
 RX MEDLINE-87231065; PubMed-3588308;
 RA Kay R.J., Russnak R.H., Jones D., Mathias C., Candido E.P.M.;
 RT "Expression of intron-containing C. elegans heat shock genes in mouse
 RT cells demonstrates divergence of 3' splice site recognition sequences
 RT between nematodes and vertebrates, and an inhibitory effect of heat
 RT shock on the mammalian splicing apparatus."
 RL Nucleic Acids Res. 15:3723-3741(1987).
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: K03273; AAA28068.1; -;
 DR EMBL: U64837; AAB04842.1; -;
 DR EMBL: U64837; AAB04839.1; -;
 DR PIR: B24289; B24289.
 DR Wormpep: T27E4.2; CE14249.
 DR Wormpep: T27E4.8; CE14249.
 DR InterPro: IPR002068; HSP20.
 DR Pfam: PF00011; HSP20; 1.
 DR PROSITE: PS01031; HSP20; 1.
 KM Heat shock; Multigene family.
 SO SEQUENCE 145 AA; 16253 MW; 06C36A1F60D15A11 CRC64;

Query Match 44.88; Score 43; DB 1; Length 145;
 Best Local Similarity 45.08; Pred. No. 4.6;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVOEIKQ 20
 Db 65 KINLDGHTLSTQGEOLKTE 84

RESULT 14
 POP6_YEAST STANDARD; PRT; 158 AA.
 AC P35218;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Ribonucleases P/MRP protein subunit POP6 (EC 3.1.26.5) (RNases P/MRP
DE 18.2 kDa subunit) (RNA processing protein POP6).
DE POP6 OR YGR030C.
OC Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT chromosome VII."
RL Yeast 13:1077-1090(1997).
CC -1- FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT
CC GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.
CC ALSO A COMPONENT OF RNASE MRP.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P
CC RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOIETY AND AT LEAST
CC 8 PROTEIN SUBUNITS, POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND
CC RPL1.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 272815; CA97018.1; -
DR SGD: S0003362; POP6.
DR Hydrolase; Nuclear protein; tRNA processing; Coiled coil.
DR KW DOMAIN 51 COILED COIL (POTENTIAL).
DR FT SEQUENCE 158 AA; 18210 MW; 6C27A73FAD521181 CRC64;
SQ
Query Match 44.8%; Score 43; DB 1; Length 158;
Best Local Similarity 64.3%; Pred. No. 5;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGV 14
DB 65 QINMADRSIGLOOV 78

```

```

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterson R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 47-143 FROM N.A.
RX MEDLINE=83220736; PubMed=6190129;
RA Russnak R.H., Jones D., Candido E.P.M.;
RT "Cloning and analysis of cDNA sequences coding for two 16 kilodalton
RT heat shock proteins (hsp) in Caenorhabditis elegans: homology with
RT the small hsp of Drosophila."
RL Nucleic Acids Res. 11:3187-3205(1983).
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M14334; AAA28070.1; ALT_SEQ.
DR DR EMBL: AC006774; AAF60616.1; -
DR EMBL: X01577; CAA25732.1; -
DR PIR: A38884; HKM41.
DR PIR: A25199; A25199.
DR Wormpep; Y46H3A.2; CE22003.
DR Interpro; IPR02068; HSP20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
DR KW Heat shock; Multigene family.
DR SEQUENCE 143 AA; 16252 MW; CID0F59D26E36C24 CRC64;
SQ
Query Match 43.8%; Score 42; DB 1; Length 143;
Best Local Similarity 45.0%; Pred. No. 6;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVQETKEQ 20
DB 69 KIKLDGRDLKIEIGIQTKE 88

```

Search completed: February 3, 2003, 13:12:15
 Job time : 6.16667 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2003, 13:07:47 : Search time 21.5 Seconds
(without alignments)
191.672 Million cell updates/sec

Title: US-09-171-432A-46
Perfect score: 96
Sequence: 1 KINLADRMGLSGVOEIKEQ 20

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
2	96	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
3	96	100.0	251	12 Q9ENQ8	Q9enq8 hepatitis a
4	96	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
5	96	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
6	96	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
7	96	100.0	251	12 Q9ENQ3	Q9enq3 hepatitis a
8	96	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
9	96	100.0	251	12 Q9ENQ0	Q9enq0 hepatitis a
10	96	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
11	96	100.0	251	12 Q9ENP8	Q9enp8 hepatitis a
12	96	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
13	96	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
14	96	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
15	96	100.0	251	12 Q9ENP4	Q9enp4 hepatitis a
16	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a

17	96	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
18	96	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
19	96	100.0	251	12 Q9ENN9	Q9enn9 hepatitis a
20	96	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
21	96	100.0	251	12 Q9ENN4	Q9enn4 hepatitis a
22	96	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
23	96	100.0	1124	12 Q84780	Q84780 hepatitis a
24	96	100.0	1161	12 Q05794	Q05794 hepatitis a
25	96	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
26	96	100.0	2218	12 Q67824	Q67824 hepatitis a
27	96	100.0	2218	12 Q67817	Q67817 hepatitis a
28	96	100.0	2225	12 Q9DL32	Q9dl32 hepatitis a
29	96	100.0	2227	12 Q67825	Q67825 hepatitis a
30	96	100.0	2227	12 Q67826	Q67826 hepatitis a
31	96	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
32	96	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
33	96	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
34	96	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
35	96	100.0	2227	12 Q9WMA9	Q9wma9 hepatitis a
36	96	100.0	2227	12 Q91FH5	Q91fh5 hepatitis a
37	96	100.0	2227	12 Q8V0N6	Q8v0n6 hepatitis a
38	96	100.0	2227	12 Q8QV03	Q8qv03 hepatitis a
39	92	95.8	251	12 Q9ENP0	Q9enp0 hepatitis a
40	92	95.8	251	12 Q9ENN7	Q9enn7 hepatitis a
41	92	95.8	251	12 Q9ENN5	Q9enn5 hepatitis a
42	91	94.8	2227	12 Q8QR16	Q8qr16 hepatitis a
43	88	91.7	251	12 Q9ENR0	Q9enr0 hepatitis a
44	88	91.7	251	12 Q9ENN8	Q9enn8 hepatitis a
45	83	86.5	251	12 Q9ENQ7	Q9enq7 hepatitis a

ALIGNMENTS

RESULT 1
Q9ENR1 PRELIMINARY: PRT: 251 AA.
AC Q9ENR1:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Al:
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -
FT NON_TER 1 1
FT 251 251
SQ SEQUENCE 251 AA: 28749 MW: 58A520D873893445 CRC64;
Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3,8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVOEIKEQ 20
DB 125 KINLADRMGLSGVOEIKEQ 144

RESULT 2
Q9ENQ9 PRELIMINARY: PRT: 251 AA.
AC Q9ENQ9:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 3
O9EN08 PRELIMINARY; PRT; 251 AA.
AC O9EN08: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A160;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047655; BAB12163.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 4
O9EN06 PRELIMINARY; PRT; 251 AA.
AC O9EN06: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A162;
RA Fujiwara K.;

RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6ED3B CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 5
O9EN05 PRELIMINARY; PRT; 251 AA.
AC O9EN05: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 6
O9EN04 PRELIMINARY; PRT; 251 AA.
AC O9EN04: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A201;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 7

OYENQ3 PRELIMINARY; PRT; 251 AA.
AC OYENQ3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A204;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047660; BAB12168.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 8

OYENQ2 PRELIMINARY; PRT; 251 AA.
AC OYENQ2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A205;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047661; BAB12169.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 9

OYENQO

ID OYENQO PRELIMINARY; PRT; 251 AA.

AC OYENQO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A302;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047663; BAB12171.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28644 MW; 1F77AA3ECF3D668B CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 10

OYENP9 PRELIMINARY; PRT; 251 AA.
AC OYENP9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A303;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047664; BAB12172.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CA5C1A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 11

OYENP8 PRELIMINARY; PRT; 251 AA.
AC OYENP8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

OX NCBI_TaxID=12092;
RN (1)
SEQUENCE FROM N.A.
RC STRAIN=A304;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047665; BAB12173.1; -
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 12

O9ENP7 PRELIMINARY; PRT; 251 AA.
ID O9ENP7;
AC O9ENP7;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28720 MW; C342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 13

O9ENP6 PRELIMINARY; PRT; 251 AA.
ID O9ENP6;
AC O9ENP6;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=A307;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047667; BAB12175.1; -
FT NON_TER 1 1

FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28644 MW; 1F77AA3ECF3D66B8 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 14

O9ENP5 PRELIMINARY; PRT; 251 AA.
ID O9ENP5;
AC O9ENP5;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179CF57A6D CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOETKEQ 20
Db 125 KINLADRMGLSGVOETKEQ 144

RESULT 15

O9ENP4 PRELIMINARY; PRT; 251 AA.
ID O9ENP4;
AC O9ENP4;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=A414;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047669; BAB12177.1; -
FT NON_TER 1 1
FT NON_TER 251 251
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOETKEQ 20

Mon Feb 3 13:52:48 2003

us-09-171-432a-46.rspt

Page 5

Db 125 KINLADRMIGSGVQEIKEQ 144

Search completed: February 3, 2003, 13:14:24
Job time : 21.5 secs

THIS PAGE BLANK (USPTO)